Whitepaper
Big Data in Earth Observation

Big Data Value Association
TF7-SG5: Earth Observation and Geospatial

November 2017

© BDVA

Big Data Value Association AISBL
BDVA c/o INDRA, Rue Froissart, 95, B-1040 BRUSSELS
Email: info@core.bdva.eu – www.bdva.eu

ACRONYMS

C3S	Copernicus Climate Change Service
CAMS	Copernicus Atmosphere Monitoring Service
CAP	Common Agricultural Policy
CBIR	Content Based Image Retrieval
CLMS	Copernicus Land Monitoring Service
CMEMS	Copernicus Marine Environment Monitoring Service
CNES	Centre National d'Études Spatiale
EASRC	European Association for Remote Sensing Companies
EC	European Commission
ECNC	European Centre for Nature Conservation
EDSA	European Data Science Academy
EEA	European Environment Agency
EMS	Copernicus emergency management service
EO	Earth Observation
EOSC	European Open Science Cloud
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GEOSS	Global Earth Observation System of Systems
GMES	Global Monitoring for Environment and Security Programme (presently Copernicus)
IOGP	International Association of Oil & Gas Producers
JRC	Joint Research Centre
LUCAS	Land use and land cover survey
NSI	National Statistics Institutes
PPP	Public Private Partnership
REPs	Regional Exploitation Platforms
SAR	Synthetic Aperture Radar
STEM	Science, Technology, Engineering And Mathematics
TEPs	Thematic Exploitation Platforms

UHI	Urban Heat Island
UN-GGIM	Global Geospatial Information Management
WMO	World Meteorological Organization

REFERENCES

A web-based platform for remote sensing image annotation. In: [Report] / auth. Garcia-Pedrero A. and Gonzalo-Martin, C. and Lillo-Saavedra, M. and Ortiz-Toto, C.. - [s.I.] : GEOBIA 2016 : Solutions and Synergies., University of Twente , Sept. 2016.

Deep neural networks for acoustic mod- eling in speech recognition: The shared views of four research groups. [Report] / auth. Hinton G., Deng, L., Yu, D., Dahl, G. E., r. Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N. and Kingsbury, B.. - [s.I.] : EEE Signal Processing Magazine 29(6) pp.82-97, 2012.

A review on automatic image annotation techniques. Pattern Recognition. [Report] / auth. Zhang D., Islam, M. M., & Lu, G.. - 2012.

A survey into the state and health of the European EO services Industry [Report] / auth. EARSC. - September 2015.

A survey into the state and health of the European EO services Industry [Report] / auth. EARSC. - [s.I.] : EARSC, 2017.

A survey of temporal data mining [Report] / auth. Srivatsan Laxman P. S. Sastry. - [s.I.] : Sadhana, 2006.

Big, Linked and Open Data: Applications in the German Aerospace Center [Report] / auth. C. Nikolaou M. Datcu et all. - [s.I.] : DLR, 2014.

—Content Based Image Retrieval at the End of the Early Years [Report] / auth. A.W.M. Smeulders M. Worring, S. Santini, A. Gupta, R. Jain. - [s.I.] : IEEE Transactions, Pattern Analysis and Machine Intelligence, 2000.

Content-based Multimedia Information Retrieval: State of the Art and Challenges [Report] / auth. Michaels Lew Nicu Sebe, Chabane Djeraba, Ramesh Jain. - [s.I.] : ACM Transactions on Multimedia Computing, Communications, and Applications, 2006.

Copernicus Market Report [Report] / auth. PwC. - [s.I.] : European Commission, 2016.

Copernicus Sentinels' Products Economic Value: A Case Study of Winter Navigation in the Baltic [Report] / auth. Geoff Sawyer Ariane Dubost, Marc de Vries, Iris van de Kerk. - [s.I.] : ESA/EARSC, 2015.

Earth observation big data for climate change research [Report] / auth. Hua-Dong Guo Li Zhang, Lan-Wei Zhu. - [s.I.] : In Advances in Climate Change Research, 2015.

Earth Observation For Oil And Gas Geo-Information Requirements [Report] / auth. Ha. - 2014.

Earth Observation For Oil And Gas Geo-Information Requirements (Deliverable 1: ESRIN/AO/1-7568/13/I-AM – Value Added Element) [Report] / auth. Hatfield Consultants Partnership In Association With: Arup, Rps Energy, C-Core, Centrum Badań Kosmicznych (Poland). - [s.I.] : ESA, July 2014.

ience for Environment Policy Future Brief: Earth Observation's Potential for the EU Environment [Report] / auth. Science Communication Unit University of the West of England, Bristol. - [s.l.] : European Commission, 2013.

Imagenet classification with deep convolutional neural networks. [Report] / auth. Krizhevsky A., Sutskever, I. and Hinton, G. E. - [s.l.] : Advances in neural information processing systems, pp. 1097–1105. , 2012.

Open data policies and satellite Earth observation [Report] / auth. Harris r. and Baumann, I.. - [s.l.] : Data policy, 2015.

Remote sensing big data computing: challenges and opportunities. Future Generation Computer Systems [Report] / auth. Ma Y., H.,Wang, L., Huang, B., Ranjan, R., Zomaya, A. and Jie, W.. - 2015.

Study to examine the socio-economic impact of Copernicus in the EU. Report on the Copernicus downstream sector and user benefits. [Report] / auth. PwC. - [s.l.] : European Commission.

Study to examine the socioeconomic impact of Copernicus in the EU. Report on the socio-economic impact of the Copernicus programme. [Report] / auth. PwC. - [s.l.] : European Commission, 2016.

TABLES OF FIGURES

1. CONTEXT AND OPPORTUNITIES

Earth Observation (EO) can be defined as the gathering of information about Earth's physical, chemical and biological systems using remote sensing technologies such as satellites and aerial sensors, supplemented (Science Communication Unit, 2013) by ground-based observations and other surveying techniques.

EO is used to monitor and assess the status of, and changes in, the natural and man-made environment. Space-based technologies deliver reliable and repeat-coverage datasets, which combined with research and development of appropriate methods, provide a unique means for gathering information concerning the planet. EO data therefore serves as an important basis for decision-making for sustainable development and disaster management. Examples include the monitoring of the state, protection and evolution of our environment, be it land, sea or air, land use, monitoring of coastlines, and the ability to rapidly assess situations during crises such as extreme weather events or during times of human conflict.

1.1 Context

Earth is facing unprecedented climatic, geomorphologic, environmental and anthropogenic changes, which require observation and monitoring at micro-, meso- and macroscales. Thus, this results in the need for a multitude of new orbital and suborbital Earth Observation (EO) sensors that revisit under mission designed frequencies their observation targets. Earth Observations embraces today a wide variety of techniques: ground penetrating radar, ground surveys, unmanned aerial vehicles, aerial remote sensing and satellite remote sensing.

Satellite based Earth Observation (EO) data volumes are increasing at a rate of several Terabytes a day: for example, Sentinel-1A, which was launched on the 3rd April 2014, already delivers high resolution SAR global data every 12 days at a rate of 2.5TB per day. Since then, other Sentinel missions became operational: Sentinel-2A and Sentinel-3A were launched on the 23rd June 2015 and the 16th February 2016 respectively. They deliver 0.8 TB (resp. 0.3 TB) of image data per day in full operational capacity. Sentinel-1B was launched on 25 April 2016, Sentinel-2B on the 7th March 2017 and recently, Sentinel 5P was launched on 13th October 2017. Other missions are soon to be launched. These will provide much larger volumes of data in heterogeneous formats (with different spatial and temporal resolutions). The promoted Copernicus free and open access policy creates unprecedented opportunities for both industrial and academic research communities.

The increasing data volume represents an immediate challenge, as does the type of observation sensors used to generate data with heterogeneous formats, semantics, measurements, resolutions, modalities and a-synchronicities. The resulting high complexity of Big Data processing is multi-faceted as it includes advanced data harmonization, pre-processing, analytics and multi-level fusion, uncertainty propagation together with the deployment of knowledge models.

Thus, volume, variety, velocity and veracity apply to this kind of datasets. Traditional data bases and processing techniques cannot cope with these new needs. Big Data science and technologies are needed for Space data exploitation in different downstream application fields.

1.2 Opportunities

According to Copernicus Market Report 2016[1], the worldwide potential EO downstream market was valued at EUR 2.8 billion in 2015. It is expected to grow to EUR 5.3 billion in 2020, at an annual growth rate of 13%. Innovations in technologies and policies offer today the opportunity to unlock this huge market potential; the key innovations are clearly Big Data Technologies and the EO Data Policy.

Big Data Technologies

The EO downstream market will benefit from the implementation of open Big Data platforms, notwithstanding the need for the deployment of re-usable Big Data analytics open services. These should ideally comply with open-geospatial standards or at least lead the way in the establishment of the next generation of open geospatial data analytics services in Europe and beyond. This will consequently change the culture for access to data but also rapidly process knowledge for real-time business operations in Europe and make them more competitive. These technologies and their impact on EO market are developed further in this paper.

EO Data Policy

The high potential of a data-driven economy and the need to strengthen the data value are already acknowledged at European level, therefore, efforts are being made regarding the sharing, use and interoperability of data, based on common standards and to build an environment ensuring appropriate protection.

The issue of managing data derived from EO satellites, in terms of access, pricing, data rights and other aspects, is commonly referred to as EO data policy. The topic covers a range of issues including: international and national laws and regulations; intellectual property rights; security; socioeconomic benefits of free and open data; public-private partnership; and pricing policy. Data policy is a combination of various concerns and interests of the operators and users.

While open data for society at large is a relatively new phenomenon, the space sector has been accustomed to open data for all of its existence, especially in astronomy and navigation, but also in Earth Observation. In this sense, Ray Harris and Ingo Baumann conducted a study (Harris, 2015) in 2015, examining 21 policy and legal instruments concerning open data made by a range of organizations related to Earth Observation data, four of them being European: European Commission, ESA, EUTMESAT, and CNES. The policy statements and legal instruments specific to Earth Observation of these four organizations are summarized in the *Annex I*.

According to the study, these open data statements and policies are predominantly in favour of full, free of charge and open access to data. The more general such statements are, the clearer they are on these principles. The closer they are to operational data supply the less they are inclined to provide complete open access to all data. This contrast may well be tied to public funding responsibility. Commenting on geospatial data in general, the United Nations Initiative on Global Geospatial Information Management (UN-GGIM) identified a tension between the expectations of geospatial data being free of charge, on the one hand, and the requirement for public funding to maintain accurate and quality-assured geospatial data on the other hand. This points to a core issue of open data: the supply of data free of charge requires continuing assured long-term public funding, demanding replacement systems for data collection in time and data archives with substantial capacity. The fast

[1] http://www.copernicus.eu/sites/default/files/library/Copernicus_Market_Report_11_2016.pdf

evolution of digital technologies is a huge challenge in this respect. Commercial re-users of data are increasingly demanding quality assurance for public data in order for them to provide reliable products and services to their clients.

The main exceptions are for foreign and national security and defense purpose, but legal reasons also play an important role. These exceptions are seldom made specific in how they are to be interpreted and implemented in practice. This leaves an open question on the real extent of open access and leaves both data providers and data users with large uncertainty with regard to individual data sets. *There is a need for much greater clarity on how exceptions will be handled in practice by the competent public authorities providing access to data*. This is especially relevant for those data bringing benefits for commercial users. *Clear guidelines are needed for dealing with privacy aspects and compliance with applicable data protection laws*.

There are unquestionable benefits in using licenses for data access, as the legal terms of use can be defined in detail. However, the use of licenses is not free of problems. Some standard licenses may not be appropriate for data, especially EO data. Licensing also sets an administrative hurdle to cross, especially for individual or scientific users. Multiple licenses in use at the national level, with different terms and conditions, may cause legal uncertainties and may discourage the commercial use of public sector data. Where multiple input data is used to create value-added products and services, the task to analyze all relevant licenses and to reflect them properly in the terms and conditions for the value-added product or services may be huge. *Development of simplified and unified license conditions applied by all public authorities of a country may greatly help in this respect*.

The current trend towards full, free and open access to public sector digital data has the potential of greatly enhancing the use of Earth Observation data on global, regional and national level, creating multiple benefits for society. However, public authorities should take care not to counter their commitments by excessive and unclear access limitations, multiple and inconsistent approaches to licenses or broad legal exclusions with regard to data availability and quality.

2. ECONOMICS & BUSINESS OF EO

The EO technologies are serving many different sectors of the economy, and bring various types of **benefits**, ranging from increased productivity up to lives saving. The positive impacts of using EO services in the value chains of the economic sectors can be translated into economic value. However, today there is no firm and agreed figures for the estimation of the benefits of EO for each targeted sector. Earth Observation is also, *per se*, an economic activity, delivering EO based services to the above-mentioned sectors by suppliers (data providers, infrastructure operators, service providers). The delivery of these services generates specific EO business **revenues**, drawn from the targeted economic sectors.

The economic importance of the EO activities, in terms of business revenues as well as benefits for several targeted sectors, has been analyzed and evaluated by various actions supported by the European Union and ESA, on one side, and by EARSC on the other side. Several market assessment studies are available, e.g. ESP report (European Space Policy Institute, Report 39, November 2011)[2] PwC market studies (Science Communication Unit, 2013)[3], (PwC, 2016)[4], (PwC)[5] and the EARSC survey (EARSC, September 2015), (EARSC, 2017). Both market research efforts have been

[2] https://www.espi.or.at/images/stories/dokumente/studies/ESP_Report_39.pdf
[3] http://www.copernicus.eu/sites/default/files/library/FutureBrief6_Feb2013.pdf
[4] http://www.copernicus.eu/sites/default/files/library/Copernicus_SocioEconomic_Impact_October_2016.pdf
[5] http://www.copernicus.eu/sites/default/files/library/Copernicus_Report_Downstream_Sector_October_2016.pdf

integrated into a comprehensive market report published by the European Union (PwC, 2016)[6]. The purpose of the present White Paper is not to duplicate or contradict the results of this huge endeavor, but rather to take advantage of the work done so far to highlight the importance of EO in the European economy and its relevance for the research on Big Data technologies.

Overall, it is assumed that the **benefits** for the main sectors in Europe will be above 2 billion Euros in 2020, with an average annual growth rate above 20%. The EC MEMO/11/469/EN (June 2011) estimates the benefits of the GMES/Copernicus services in EU up to 130 billion Euros over the period 2006-2030, i.e. around 6.9 billion Euros per year. According to the EARSC survey published in 2015[7] ("A survey into the state and health of the European EO services industry"), the total revenue for the EO services industry in 2014 was 910 million Euros, with an estimated growth rate above 7%.

The EARSC survey published in 2017 reports revenue of 1145 million Euros for 2016, confirming the growth rate above 10%. According to EARSC report, 45% of the revenues originate from outside Europe (USA, Africa and Middle-East, Asia, Latin America, Russia). The current figures (year 2016) do not yet reflect the positive impact on the market that Sentinel data will have on the availability of improved or new services, and consequently the revenues for the EO sector.

So far, most of the EO data and services are "consumed" by the Public sector; for example, the US Government spending represents 85% of EO imagery market (DigitalGlobe derives nearly 50% of its revenues from the EnhancedView contract with the National Geospatial Intelligence Agency).

In Europe, there is a significant level of public spending related to the use of EO data by research and Academic institutions; the corresponding total yearly costs are not taken into account in the EARSC nor PwC reports; however, they could represent a multiple of the revenues identified by EARSC.

The report published in October 2016 by PwC (PwC) analyses several key sectors potentially impacted by EO services. The table hereafter highlights, per sector, the estimated revenues for the EO services industry in 2015 and 2020 for Europe.

Sector	Revenues for EO in 2015	Revenues for EO in 2020	Comments
Agriculture	71M€	221 M€	Focus on precision farming only
Forestry	36 M€	41M€	
Renewable Energy	22 M€	44 M€	Solar, wind, hydro, biomass
Urban development	45 M€	99 M€	Urban monitoring
Ocean monitoring	103 M€	116 M€	
Oil & Gas	73 M€	--	Consistent estimates for 2020 are not provided

Table 1 Estimated revenues for EO services industry

The expected benefits and revenues can be realized on two main conditions:

[6] http://www.copernicus.eu/sites/default/files/library/Copernicus_Market_Report_11_2016.pdf
[7] http://earsc.org/news/earsc-eo-industry-survey

1. The EO community must implement the appropriate business models to transform the EO activities into a sustainable, self-supporting business addressing the demand of the users in the various economic sectors in the most efficient and cost-effective way.
2. The EO and ICT technologies together must manage to make available suitable solutions to address the Big Data challenges of exploiting EO data. This second point is the main focus of this White Paper.

These conditions represent challenges and call for innovation in the ICT sector as well as in the implementation of new market and business concepts.

3. EARTH OBSERVATION DATA APPLICATION AREAS

The topics described below do not represent the full breadth of activities related to the exploitation of Earth Observation (EO) data, but it is hoped that they give an indication of this ever-growing and advancing field of work.

Several economic sectors are strongly impacted by EO domain:

- Agriculture
- Forestry
- Safety / Security / Disaster Monitoring / Risk Assessment/ Insurance
- Energy / Oil and Gas/ Resource Management
- Ocean Monitoring: Marine and Maritime / coastal zone management/ Fisheries and Aquaculture
- Public Health
- Urban Development/ Smart cities/ Cultural Heritage
- Transport - Navigation

Each of these domains and the impact of EO data, applications and services are described in *Annex II*.

4. EUROPEAN JOINT EFFORTS ON BIG DATA

Despite existing technological difficulties, opportunities for business development arise. Moreover, space open data policies (providing quality information to end users) or Big Data infrastructure development / maintenance are of high interest at European level. Therefore, policy makers focus on legislation frameworks, and a Public Private Partnership is settled in support of these opportunities.

Several EU important initiatives developed within the EU R&I Framework Programmes paid attention to Big Data and data mining aspects. The most relevant are the ones that had as a result the implementation of the Copernicus core services: Atmosphere (CAMS), Marine (CMEMS), Land (CLMS), Climate (C3S), Emergency (EMS) and Security. These are briefly described in *Annex III - Copernicus Core Services*.

Given the importance of the Big Data aspects in space domain, starting June 2015, a new Unit - Space data for Societal Challenges and Growth - has been established within DG Internal Market, Industry, Entrepreneurship and SMEs / Space Policy, Copernicus and Defense. Direct contact of Big

Data Value Association[8] (BDVA) with the unit, through a dedicated working group for Earth Observation from the Association, would foster communication with interested industry.

The importance that the Commission shows to Big Data from Space is also proved by the fact that several research calls for proposals have comprised related topics within the EC's Framework Programs for Research and Technological Development. In particular, within Horizon 2020 (the current Framework Program), EO activities are considered an essential element to accompany the investments made by EU in Copernicus - the European EO and monitoring program.

In turn, ESA's view on the need for EO exploitation platforms networking has also been focusing on different efforts to incentivise the market.

4.1 EC and ESA Projects

H2020 has run two work programs since 2014, 2014-2015 and 2016-2017, and is now launching the 2018-2020 calls. Specifically, within the LEIT-SPACE work program there are specific calls dedicated to Earth Observation in all periods, focusing on the evolution of Copernicus and the exploitation of existing European space infrastructure by promoting the development of innovative products and services based on remote sensing, geo-positioning or other types of satellite enabled data. Information regarding the funded, ongoing and forthcoming H2020 projects related to EO can be found in *Annex V – EC and ESA EO big data projects*.

Leaving out the Earth Observation calls included in the *Annex V* – EC and ESA EO big data projects, there are additional topics addressing Earth Observation related technologies in other H2020 areas such as Societal Challenges. In particular, Earth Observation technologies are of particular interest in Societal Challenge for climate action, environment, resource efficiency and raw materials. In this area, the focus is on GEOSS, notably the development of comprehensive and sustained global observation and information systems[9]. Other areas where Earth Observation is relevant are Blue Growth (Ocean observation technologies), Sustainable Food Security (smart farming), both included within the Societal Challenge 2, Competitiveness of the European Space Sector (COMPET), among others.

In addition to these calls, another European intergovernmental organization, ESA, closely supports the same topics. ESA is Europe's gateway to space, having as main mission to 'shape the development of Europe's space capability and ensure that investment in space continues to deliver benefits to the citizens of Europe and the world (www.esa.int). Given the fact that the EO data volumes are increasing due to the new ESA launched Sentinel missions within the European Copernicus programme, more and more initiatives regarding Big Data are funded under different ESA programmes, and the topic is far from being completely addressed, opening important opportunities both for the industry dealing with data mining / Big Data technologies and for the EO service providers that benefit from the respective technologies.

[8] http://www.bdva.eu

[9] Example of previous topics: SC5-18-2017 - Novel in-situ observation systems; SC5-19-2017 - Coordination of citizens' observatories initiatives; and SC5-20-2016 - European data hub of the GEOSS information system

A list with implemented and ongoing projects funded either by EC or ESA is to be found in *Annex V – EC and ESA EO big data projects*.

4.2 Collaborative Ground Segment

The Sentinel Collaborative Ground Segment complements the exploitation of the Sentinel missions in a number of areas, such as:

- Data acquisition and (near-) real-time production. Local ground stations are configured to receive Sentinel data as the satellite passes overhead;
- Complementary products and algorithm definitions – these 'collaborative data products' may be tailored for regional coverage or specific applications. These types of products may extend the Sentinel core product chains;
- Data dissemination and access, supporting redistribution of Sentinel core products by establishing additional pick-up points;
- Development of innovative tools and applications;
- Complementary support to calibration/validation activities.

As well as the challenging task of building and launching a satellite, the success of an Earth Observation mission relies on the ability to operate the satellite from the ground and ensure that the data gathered are of good quality and readily made available to users.

To achieve this, a control center with an antenna to send commands to the satellite is needed, together with at least one ground station to receive the data as they are 'downlinked' from the satellite as it passes overhead. Facilities for data processing, storage and distribution are also needed.

Last but not least, the quality of the data and the performance of the satellite sensors have to be continuously monitored. If needed, actions such as 'calibration activities' have to be promptly carried out, guaranteeing the overall performance of the mission. These functions are all ensured by the ground segment.

Each Copernicus satellite mission, both the dedicated Sentinel missions as well as each contributing mission, has a ground segment, and is independently operated.

While each ground segment is independent, they are all linked to form the Copernicus Ground Segment. This large space capacity is coordinated through the Space Component Data Access System. Here, data flows from the various ground segments are streamlined – in a way that is as transparent as possible to the users.

The Copernicus Ground Segment is complemented by the Sentinel Collaborative Ground Segment which was introduced with the aim of exploiting the Sentinel missions even further. This entails additional elements for specialized solutions in different technological areas such as data acquisition, complementary production and dissemination, innovative tools and applications, and complementary support to calibration and validation activities.

The overall space capacity, beyond the single missions, is coordinated through the Copernicus Space Component Data Access System.

The system provides comprehensive and coordinated access to space data, to:

- transparently link the different data providers and the various Copernicus services using specific coordinating functions;
- create synergy and sustainability across the various contributing missions;
- access a simplified interface for advertisements and the service desk rather than using multiple data provider interfaces.

The system is the hub of an interoperable network of distributed European ground segments contributing to Copernicus, culminating in a harmonized, one-stop-shop for users.

Data and services are accessible in the form of datasets, which are pre-defined collections of coherent (single and/or multi-mission) products that are derived from service requirements after trade-off considering the overall capacity of the space component.

The Sentinel Core Ground Segment and the Data Access Coordinated System are directly managed by ESA while other parts are managed by third parties such as National Space Agencies and interface to the core ESA elements via specific agreements with the Agency.

4.3 Thematic Exploitation Platforms

Thematic Exploitation Platforms (TEPs) are ESA's answer to the technical challenges related to the ever-increasing volume of EO data. The TEP concept aims to supply a virtual working environment which provides users with access to data, algorithms and computing resources and tools, thus avoiding the need to download, store and process large volumes of data. A range of TEPs are being in pre-operational phase, each tailored to a theme for a specific user community and research field (e.g. Hydrology TEP, Polar TEP, Coastal TEP). These innovative virtual platforms aim to simplify the extraction of information from EO data, for the advancement of data-intensive research within specific areas of environmental research. This new form of providing data access and processing capabilities will encourage the wider exploitation of EO data. The current TEPs are briefly described in *Annex IV - ESA Thematic Exploitation Platforms.*

4.4 Regional Exploitation Platforms

The Regional Exploitation Platforms (REPs) represent a step forward and are based on the same concept as the TEPs, bringing together: data centers (EO and non-EO data), computing resources and hosted processing, collaborative processing tools, application shops, communication tools (social network) and documentation. Unlike the TEPs that are focusing on specific user communities (domains), the REPs are addressing regional areas (e.g. West Africa, South-East Europe, etc.).

4.5 Data Access and Information services - DIAS

As the huge potential of data and information produced by Copernicus became clear, the key to unleash it is the easy access to it. In this regard, the Commission created the background for establishing an Integrated Ground System (IGS) for Copernicus that will enhance the user communities to have the maximum benefits from EO data and information. The Commission is keen to ensure that all Member States and Copernicus participating countries, on equal basis, can take advantages on the Big Data paradigm through the DIAS (Data Access and Information Services) initiative.

The objective of DIAS is to enable the creation of third party advanced services exploiting EO data and information for the benefit of end-users. DIAS complements the existing Data Dissemination System and intends to "bring the users to the data".

DIAS will functionally consist of 3 types of services:

- **Back office services** - will contain all Copernicus Data and Information together. They are centered on making available Copernicus data and information in a scalable computing environment fit for their exploitation;
- **Integration services** - consist mainly of interface layers and have a key role to play in the different interoperability requirements; they offer an environment to create, build and exploit applications. This environment is to be developed and managed by the DIAS Service Providers according to their business model(s);
- **Front office services** are centered on the provision of services by third parties based on the exploitation of the Copernicus data and information made available through the back office services.

The success of DIAS highly depends not only on the strong relations between Copernicus actors but also on the involvement of Member States and participating countries, Information and Communication Technology (ICT) and Earth Observation (EO) industries as well any third party interested in using Copernicus data and information.

5. BIG DATA IN SPACE – TECHNOLOGIES

In the last years, different data mining technologies were developed in order to cope with the different volume, variety, velocity and veracity of space based data, in order to make the EO services and applications development more efficient and to most benefit from all the information hidden within the data. Relevant Big Data technologies are being presently applied in EO domain but still challenges remain which are briefly described below according to the different technology areas of the BDV Reference Model:

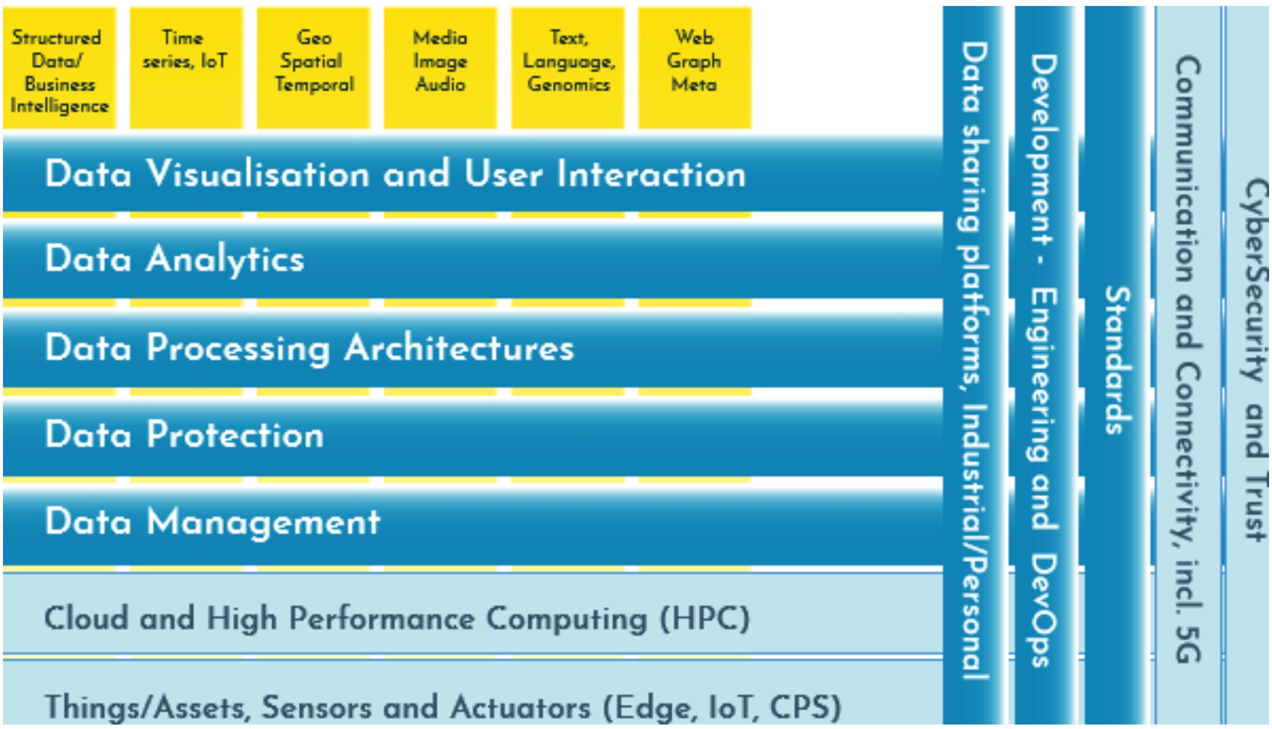

Figure 1: Big Data Value Reference Model

The BDVA Big Data Value Reference Model (from the SRIA 4.0[10]) is shown in the figure above.

The following describes Big Data in Space technologies according to these areas in the model:

- Data Visualization and User Interaction;
- Data Analytics;
- Data Processing Architectures;
- Data Management: - including Big Data Types and semantics with a focus on Geospatial and Image data types;
- Cloud and High Performance Computing (HPC) private/public/hybrid Infrastructures.

5.1 Data Visualization

Visual Data Mining - VDM

VDM is a generic term to refer to visual data exploration tools, which allows interactive data presentation in order to increase users' capabilities to understand the information content of large data sets of images and extract meaningful, relevant semantic clusters, together with quantitative measurements presented in a suggestive, visual way.

The VDM tools are able to offer a preliminary insight into a set of optical or radar data, by revealing its semantic structure (natural groups/clusters of resembling scenes/tiles) and some quantitative estimations regarding that structure, through suggestive, simple visual representations. Based on VDM maps, charts and diagrams, the user can make better and faster a priori estimations on the feasibility of the desired image processing.

3D visualization

3D visualization techniques provide capabilities for visual analytics of geospatial and time series data and building Augmented Reality applications. 3D provides a web based, interactive visualization tool for geospatial data and related information, using state of the art technologies enabling flexible and fast interactive visualization of big multidimensional spatial data through linked views. The services can combine the geo-visualization with general non-geo approaches. Often technologies are used to combine the wealth of data about conditions on Earth with the technology to display that information on a globe. New technologies are developed to visualize wind, pressure, wave height, thermal, UV levels, or other Earth observation data on a 3D globe right in the web browser.

5.2 Data Analytics

Machine and deep learning

The daily generation of terabytes of Earth Observation data opens opportunities to both industry and scientific communities to fully exploit these data in order to provide services to end-users (Ma, 2015). However, the analysis of these voluminous data is not possible by manual means, and therefore the use of automatic and semiautomatic tools is required. In this regard, several machine learning and

10 http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.pdf

pattern recognition techniques have been used in remote sensing (RS) applications in order to facilitate the image processing tasks.

Recently, deep learning-based methods have demonstrated excellent performance on different artificial intelligence tasks, including speech recognition (Hinton et al., 2012a), natural language processing (Hinton, 2012), and computer vision (Krizhevsky, 2012). As far as our knowledge goes, the use of Deep Learning for remotely sensed imagery processing is relatively recent. Particularly, it has been used in remote sensing area for image preprocessing, pixel-based classification, and target recognition and to the recent challenging tasks of high-level semantic feature extraction and remote sensed imagery understanding.

These methods require a training step, during which most methods require a large amount of annotated (labeled) samples that provides a good statistical differentiation of the classes to provide final accurate results. The generation of that labeled (annotated) collection is a challenging task, especially for large areas mapping, this being one of the main challenges to be solved in order to completely exploit the values inside EO data through these learning approaches. In this regard, some collaborative platforms are currently being developed with the aim of providing that annotated information (Garcia-Pedrero, Sept. 2016).

5.3 Data Processing

Complex Event Processes

CEP is a technique in which incoming data about what is happening (event data) is processed in real time to generate higher-level, more-useful, summary information (complex events). Event processing platforms have built-in capabilities for filtering the incoming data, storing windows of event data, computing aggregates and detecting patterns. In a more formal terminology, CEP software is any computer program that can generate, read, discard and perform calculations on events. A complex event is an abstraction of one or more raw or input events. Complex events may signify threats or opportunities that require a response from the business. One complex event may be the result of calculations performed on a few or on millions of events from one or more event sources. A situation may be triggered by the observation of a single raw event, but is more typically obtained by detecting a pattern over the flow of events. Event processing deals with these functions: get events from sources (event producers), route these events, filter them, normalize or otherwise transform them, aggregate them, detect patterns over multiple events, and transfer them as alerts to a human or as a trigger to an autonomous adaptation system (event consumers). An application or a complete definition set made up of these functions is also known as an Event Processing Network (EPN).

5.4 Data Management – including Image and Geospatial data types

Satellite Image Time Series - SITS

Considering the continuous, yet, irregular, data acquisition, the opportunity to generate and evaluate Satellite Image Time Series (SITS) became necessary for extracting significant information regarding the complex transformation and evolution processes on the ground. These techniques providing solutions for the automatic discovery of regularities, relationships and especially temporal interdependencies are leading to a better and easier understanding of the underlying processes that causes the detected changes (Srivatsan Laxman, 2006).

The need to access and interpret SITS becomes more important with the growing interest for identifying, monitoring and analyzing dynamic and precise spatiotemporal structures like cities, agricultural fields and forests or natural phenome (floods, landslides). The observation of land cover dynamics requires data time series spanning over 20 years and more, as major environmental evolution may last longer than seasonal changes. Various phenomena interfere with day by day life aspects, such as climate changes, relief transformations or water level variations. They are all interconnected and influence the population spread and living. On the other hand, some of the human activities have strong impact on the environmental metamorphosis. Global warming or landscape modelling due to human construction activities are just some examples. Consequently, the study of Earth's surface dynamics is an important issue in the field of remote sensing imagery processing. Monitoring the land cover evolution is essential for understanding the environmental changes. This helps to distinguish between sudden variations (over short period of time, i.e. days or weeks) and gradual transformations (over long period of time, i.e. seasons - years).

Presently, there are huge amounts of data, suitable for SITS data mining, covering more than 20 years, with new data added every day. The new generation of high resolution sensors, giving access to detailed image structures, should be better exploited.

The current functionalities of the EO Payload Data Ground Segment (PDGS) systems have to be enlarged with extraction and access to the information content of SITS, for enhanced and long-term data exploitation of already acquired and forthcoming data (e.g. SENTINEL 2 data).

There is a broad range of applications involving the exploitation of the rich information content of SITS. Several examples are listed below:

- Urban dynamic monitoring, urban management and regulation;
- Critical infrastructure monitoring;
- Identification of land usage categories in time: classification of agricultural fields/forest/ populated areas;
- Wet areas monitoring: coastal zones and waterline evolution, coastal morphology, islands evolution;
- Landslides identification. Understanding the process's dynamic. Landslide morphology assessment;
- Mining areas monitoring;
- Identification and monitoring of areas affected by various phenomena or variation of vegetation;
- Disaster management;
- Land displacement monitoring (following earthquakes or close to strategic infrastructure);
- Glacier displacement monitoring;
- Health analysis for ecosystems;
- Monitoring of dynamic processes in mountain ravine;
- Monitoring of diseases evolution;
- Water management for its sustainable use.

Content Based Image Retrieval - CBIR

The main idea in content based retrieval consists of searching semantic similarities. However, the analysis of the collection by a machine is only able to provide similarity by data processing, resulting in a central issue because the meaning of an image is rarely self-evident. This can define a semantic difference between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation (A.W.M. Smeulders, 2000). This is why, in the last decade, significant interest has been shown bringing the semantic gap between the "information" in terms of low-level features extracted from an image and the "knowledge" specific to each user and application. Building ontology to express a semantic meaning of images is considered to be a really challenging issue repeatedly approached in the field of image retrieval and automatic semantic annotation (Michaels Lew, 2006).

Within a CBIR system, several steps are usually followed. Primitive features are extracted from the images in the database and indexed using classification and segmentation methods based on similarities measurements. The features inside a query image selected by the user are compared with the entire characteristics space. The system returns to the user the images mostly alike to the query, enabling the user to refine the search (A.W.M. Smeulders, 2000). Currently, new approaches are arisen for performing semantic indexing based on Machine Learning (Zhang, 2012).

6. BIG DATA – INFRASTRUCTURE

The need to move geospatial data analysis, and more specifically EO data processing, into the "cloud" has been recognized by many organizations worldwide. Consequently, several organizations and initiatives worldwide have already started, or are preparing for the uptake of EO data into their Big Data infrastructures. Still, no EU champion has been raised on this domain.

6.1 Public infrastructure

Public sector initiatives with the aim to establish Big Data infrastructures capable of processing Sentinel data have mostly sprung up from national and international space agencies already involved in setting up the Sentinel Ground Segment.

At national level, progress towards the establishment of collaborative infrastructures for processing Sentinel data has e.g. been made by some European countries.

Each country has followed its own approach, with different levels of involvement of the private industry. For example, the French THEIA Land Data Center[11] is a national inter-agency organization designed to foster the use of images coming from the space observation of land surfaces. The Centre organizes the French skills in this field and relies on a distributed data center infrastructure in order to provide data to more than 400 laboratories and 100 graduate schools on the national territory. In the United Kingdom, the Climate, Environment and Monitoring from Space (CEMS) facility (also referred to as Satellite Applications Catapult[12] was established. It is a purpose-built facility offering users access to CEMS data and various platform services. INTA (National Institute in Aerospatial Techniques[13]) is, in Spain, together with other Spanish Copernicus relays the organization that promotes the use of Copernicus Programme data, both from its Services and Sentinel images points of view.

On an international level, EGI[14] has already integrated the Geohazard Exploitation Platform (GEP) with the EGI infrastructure and has established different pilots to integrate a second ESA thematic exploitation platform (TEP), Hydrology, with the objective to increase exploitation of EO data by public and commercial organizations through the EGI e-Infrastructure. EGI is particularly interested in making leverage of the EGI's e-infrastructure offering for hosting the R&D phase, integrating EO space data with other key scientific datasets into a multi-disciplinary environment. EGI intends to establish a new way of exploiting EO data together with key scientific datasets, through services that supply the computing and storage resources needed for data access and exploitation and provide the

[11] http://www.theia-land.fr/
[12] https://sa.catapult.org.uk/cems#
[13] http://www.inta.es/opencms/export/sites/default/INTA/en/
[14] www.egi.eu

tools to manage the datasets in a distributed environment. Last, but not least, EGI aims also at providing a European-wide long-term storage opportunity for Copernicus data ensuring its high-availability for Researchers and for serving SMEs / Industries operating the DIAS platforms and marketplace guaranteeing their long-term availability and usability and realizing economies of scale.

6.2 Private infrastructures

Many commercial IT companies have started to offer cloud services for processing of EO data.

The most advanced offer – and for scientific purposes free –comes from Google through its Earth Engine[15]. Quoting their website "Google Earth Engine brings together the world's satellite imagery and makes it available online with tools for scientists, independent researchers, and nations to mine this massive warehouse of data to detect changes, map trends and quantify differences on the Earth's surface." Google plans to host all Sentinel data, putting the Earth Engine in a good position to become a premier platform for the scientific exploitation of Sentinel data.

Another commercial IT giant that has recently moved into the earth observation market is Amazon. Amazon offers through its Amazon Web Service (AWS) access to Landsat 8 data from 2015 onwards[16]. The data themselves are free, but users need to pay for Amazon cloud computing services, whereas costs are to cover the following three components: compute, storage, and data transfer out. Huang et al. (2010) describe their first experiences when implementing a geospatial data and service catalogue on AWS.

European based commercial IT providers have also identified Earth Observation as a potential business case for their cloud service offers. Within the framework of the Helix Nebula (HN) partnership (http://www.helix-nebula.eu/), often referred to as the European Science Cloud, they have engaged in a cooperation with three of Europe's leading research organizations – including ESA – in order to provide computing capacity and services that elastically meet big science's growing demand for computing power. So far, the initiative has deployed and validated three high-profile flagship experiments in high energy physics, life sciences and earth science, on commercial cloud services hosted by multiple suppliers.

All of the above examples involve commercial cloud providers that have moved into the field of earth observation, but the other way is also of course possible. Just as one example for such a venture, is the Munich-based company CloudEO (www.cloudeo-ag.com).

6.3 Towards Hybrid Infrastructures

The European Open Science Cloud (EOSC[17]) is a vision for a federated, globally accessible, multidisciplinary environment where researchers, innovators, companies and citizens can publish, find, use and reuse each other's data, tools, publications and other outputs for research, innovation and educational purposes. Making this vision a reality is essential to empower Europeans to tackle the global challenges ahead. The EOSC is one of five broad policy action lines of the European open science agenda endorsed also by the EC Communications on the Digital Single Market (DSM) strategy.

[15] https://earthengine.google.org
[16] http://aws.amazon.com/public-data-sets/landsat/
[17] https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

The European Open Science Cloud will empower research data sharing, data stewardship and data reuse in Europe for the benefit of innovation and growth.

Many of the resources and services needed for the Open Science Cloud already exist; *while technical challenges remain, most of the barriers are ones of policy and concern funding, lack of interoperability, access policies and coordinated provisioning.* The Open Science Cloud will address these issues and enrich and further advance the portfolio of resources and services to make the entire scientific lifecycle more open and transparent.

A first phase in the development aims at establishing a governance and business model that sets the rules for the use of EOSC, creating a cross-border and multi-disciplinary open innovation environment for research data, knowledge and services, and ultimately establishing global standards for interoperability for scientific data[18]. The recently approved EOSC-hub project[19] will create the integration and management system (the Hub) of the future EOSC that delivers a catalogue of services, software and data from major European research e-Infrastructures.

EOSC and DIAS

The European Commission has created and enabled an EU Earth Observation (EO) Data ecosystem for research and business building on a partnership between the EC and the European Space Agency. In this context, the EC is procuring for Copernicus consolidated Data Access and Information Services (DIAS) platforms primarily addressing open data and other components from the EO Community that will eventually lead to (data-related) EO commercial services.

The commercial services based on DIAS would become available to the scientific community also through the EOSC.

These services provided by the private sector will have to be integrated in the EOSC via procurement. Public e-Infrastructures would act as aggregators of demand for such commercial services, purchase services and distribute the acquired resources through the EOSC hub for the benefit of the scientific community.

As result, the integration of private services in the EOSC will:

- Set-up a value-added procurement mechanism for commercial services that creates economies of scale in procuring and brokering resources for the EOSC and helps optimizing investments enabling long-term sustainability for the EOSC;
- Enable the creation of new market opportunities and new solutions by facilitating cross-fertilization of heterogeneous data and services;
- Widen the offer of available services for researchers in the EOSC;
- Widen the Copernicus service offering for scientific constituencies;
- Lower access barriers for scientific community to commercial services, including DIAS supported added value services and enhance industry's potential to take advantage of scientific market opportunities.

[18] EOSCpilot project: http://www.eoscpilot.eu/
[19] EOSC-hub project: http://go.egi.eu/eosc-hub

7. EO BIG DATA – SOCIAL DIMENSION

The Earth, from its interior to the land surface, rivers, lakes, oceans and atmosphere, is a dynamic system which is continuously changing by nature as well as due to human activities. A huge amount of observations is demanded to generate valuable information regarding the state and evolution of the planet. The benefits from improving EO technologies in this scope have been the subject of decades of research and discussion. The need for understanding environmental dynamics is even more urgent nowadays, given the recognized issues of climate change, sustainable food sources and increased need for energy. In this sense, a huge number of sensors to be boarded on artificial satellites have been developed from different scientific and technical endeavors. These sensors provide observations that address a broad range of societal needs. However, earth observation systems must be highly efficient to meet these needs within challenging budgetary constraints. Thus, to fully exploit the large, expensive monitoring systems, such as EO satellites, or in-situ monitoring networks, there is a need for applications and identifiable near-term uses that brings benefit to society. Simultaneously, it is necessary that innovative and pragmatic uses of EO, to be proven in supporting policy, business, and management decisions of public and private organizations, is needed. Some examples of the general areas where EO can have a high impact in the society are:

- **Weather and climate forecasts** - European agencies have now the capability of 3-to 10-day weather forecasts, being competitive with USA capabilities. However, the improving of current atmospheric models requires new observing tools and enhanced computing capacity.
- **Monitoring our water and food resources** - one of the main resource for any country is the water supply. The information regarding its availability, quality and quantity of this resource is of paramount importance for economic development, public health and safety, food production, and recreation.
- **Ensuring healthy, productive oceans** – thanks to the measurements registered by remote sensors, the understanding of oceans and their interactions with the atmosphere have passed a revolution. Nevertheless, this is not enough to mitigate the new risks provoked by the climate change, like the ocean warming, changes in circulation, or acidification.
- **Vulnerable ecosystems monitoring** - Earth’s surface is suffering dramatic changes that have been quantified with the support of EO data. In order to keep the vulnerable ecosystems safe, many challenges must be solved, and more data are needed to tackle them.
- **Contributing to human health -** the last remote sensor generation has provided highlighted advances in pathogen, vector, and reservoir and host ecology detection, facilitating the evaluation of a greater range of environmental factors allowing prediction or early detection of disease transmission, vector production, and the emergence and maintenance of disease foci, among other aspect directly related with human health.

The presented examples highlight the social impact of EO big data. Specifically, its economic aspects, since its development could increase food production, decrease health service costs or minimize natural disasters effects.

Additionally, other types of benefits should also be taken into account, such as, the number of lives saved, improvement in the quality of the environment, the planet health. To assess all these benefits, the implication of all stakeholders is needed at different levels.

Finally, EO technologies should support policy makers in reaching the Sustainable Development Goals agenda of the United Nation.

8. BIG DATA – EDUCATION

In order to fully exploit the potential of Big Data, a key challenge for Europe is *to ensure the availability of highly skilled people*. **Data Scientists** are needed, with solid knowledge in statistical foundations and advanced data analysis methods and a thorough understanding of scalable data management, with the associated technical and implementation aspects. **Data Engineers** need to be formed to develop and exploit techniques, processes, tools and methods for applications that turn data into value; to understand the domain and the business of the organizations; to bring knowledge and connect the technology with the application domains and business.

In this scenario, some initiatives can be found at European level. The European Data Science Academy (EDSA[20]) designs curricula for data science training and data science education across the European Union (EU). EDSA establishes a virtuous learning production cycle whereby they: a) **analyze the required sector specific skillsets for data scientists** across the main industrial sectors in Europe; b) develop **modular and adaptable data science curricula** to meet industry expectations; and c) deliver **data science training** supported by **multi-platform** and **multilingual** learning resources. The curricula and learning resources are continuously **evaluated** by pedagogical and data science experts during both development and deployment. EDSA is monitoring trends across the EU to assess the demands for particular Data Science skills and expertise. Using interviews with Data Science practitioners, an industry advisory board representing a mix of sectors and automated tools for extracting data about job posts and news articles, they are building dashboards to present the current state of the European Data Science landscape, with the data feeding into our curricula development. Amongst the partners of this project are The Open University, Open Data Institute, KTH, IJS, University of Southampton, Eindhoven University of Technology (Tue), Fraunhofer Institute.

Also, the Eit Digital Master School[21] can be mentioned. The newly established Data Science Master's offers a unique academic programme, whereby students can study data science, innovation, and entrepreneurship at leading European universities. In this programme, students will learn about scalable data collection techniques, data analysis methods, and a suite of tools and technologies that address data capture, processing, storage, transfer, analysis, and visualization, and related concepts (e.g., data access, pricing data, and data privacy). Several prestigious European Universities are involved in this project: Eindhoven University of Technology (TU/e), Universidad Politécnica de Madrid (UPM), Universite Nice Sophia Antipolis (UNS), and Polytechnic University of Milan (Polimi). Most of the specializations that the students can achieve are strongly related with Big Data issues:

- Infrastructures for Large Scale Data Management and Analysis at UPM;
- Multimedia and Web Science for Big Data at UNS;
- Business Process Intelligence at TU/e;
- Distributed Systems and Data Mining for Really Big Data at KTH;
- Design, Implementation, and Usage of Data Science Instruments at TUB;
- Machine Learning, Big Data Management, and Business Analytics at Aalto University.

Moreover, EC has funded some projects aiming to establish the data scientist as a profession. One of the most representative is the Horizon 2020 **Edison project**[22]**.** This will be achieved by aligning industry needs with available career paths, and supporting academies in reviewing their curricula with respect to expected profiles, required expertise and professional certification. This will result in a

[20] http://edsa-project.eu
[21] https://masterschool.eitdigital.eu/programmes/dsc/)
[22] http://edison-project.eu/edison/edison-project

significant increase in the number and quality of data scientists graduating from universities and being trained by other professional education and training institutions in Europe.

The **BDVe**[23] project (Big Data Value ecosystem) should be also mentioned. Its mission is to support the Big Data Value PPP in realizing a vibrant data-driven EU economy and support the implementation of the Big Data PPP to be a success. In particular, the project dedicates specific activities to the objective of developing skills, education, and Centers of Excellence around Big Data that facilitate European-level coordination, it helps to align curricula with industry needs and it supports skills development thus increasing the number of European data scientists by 2020.

Some initiatives can be found outside the EU:

- The University of California, Riverside (UCR) and NASA's Jet Propulsion Laboratory has developed the FIELDS[24] program to train the next generation of scientists and engineers in large-scale data analysis and visualization. In particular, the FIELDS Summer Internship at NASA's Jet Propulsion Laboratory provides the opportunity for UCR undergraduates with majors in a STEM fields to perform research under a UCR faculty member. Successful careers in STEM depend on having the skills to understand, analyze and extract information from large amounts of data.
- The University of Penn State[25] is working in a global initiative (Building a Big Data Analytics Workforce in iSchools[26]) with the objective to develop three innovative learning approaches. These are founded on both group-based and contextualized learning methods, applicable and accessible to students majoring in disciplines outside of computer science but related (i.e., iSchools). In the frame of it, two specific projects related with Geospatial data are included: (i) **Geospatial Big Data Analytics and Geographic Education**, aiming to investigate the current situations on higher education on geospatial (big) data science in the U.S. and discuss the future direction of it. (ii) **Geospatial Big Data Visualization with Technology,** aiming to investigate the current trends in geospatial (big) data visualization using various cutting-edge technologies including high performance computing and/or immersive virtual environment and their domain applications.

Even though there are some master programs in the geospatial field that include data science skills, it is of paramount importance to perform a deeper search regarding the specific skills required by a data scientist for working in the geospatial area. In the specification of these skills, academy and industry should work together, with the aim of specifically designing careers that address the European demand for professionals and researchers with knowledge in big geospatial data exploitation.

9. CONCLUSIONS

9.1 Challenges

The need for much more efficient and faster Big Data exploitation, with timely delivery of meaningful extracted information represents a major current target in the EO domain, and therefore a major challenge to overcome in the years to come. Furthermore, EO satellite data are to be integrated with other types of data for delivering value added services and applications for Europe. The first and most

[23] http://cordis.europa.eu/project/rcn/206401_es.html
[24] http://bigdata.ucr.edu/fields/internships/summer-interships/
[25] http://sites.psu.edu
[26] https://sites.psu.edu/bigdata/

used so far are **in situ data**, which complement / validate the data gathered from space-borne and air-borne remote sensing.

With the latest development in the concept of the **Internet of Things (IoT)**, data coming from an even larger number of sensors can be successfully integrated with EO services and applications. Unstructured **data from social media networks** are also becoming more and more prominent and attractive to be integrated in new application services for various domains. This makes their integration under the Big Data realm even more challenging but with the aim of extracting knowledge at multi-geospatial scales, in real-time, at greater context and high situation awareness for critical decision-support. Even further, **citizen science** is gaining more attention from the research community, numerous initiatives (e.g. MYGEOSS project by the European Commission to develop GEOSS-based (Global Earth Observation System of Systems) smart Internet applications informing European citizens on the changes affecting their local environment) include this kind of collaboration mostly for collection, but also for analysis of data (including Crowdsourced). With respect to the EO application domains, the possibility of its integration should be considered.

Challenge-1: Generic framework for multi-level Big Data processing

Over the years, many flagship projects addressed the processing of large volume of data for critical decision-support under multiple domains. They also addressed the use of geospatial and geoscientific standards for the description of meta-data and data, including sensor observations for web processing and mapping services. Nevertheless, the standardization of further specialized data fusion, analytics and adaptive modelling with uncertainty control services remains challenging. Notwithstanding the challenge for estimating the propagation of uncertainty across the multiple Big Data sources at pre-processing and post-processing levels. With the ever-increasing complexity in Big Data dimensions, this challenge keeps increasing at an unprecedented rate. As a result, the adoption of a common framework for multi-level Big Data processing is required in order to contain the challenge. Consequently, many science and engineering communities will need to join forces to focus on and address the sub-challenges under this generic level of intelligent data processing. The concept of multi-level JDL (Joint Director of Laboratories) data fusion for decision-support is one of the best candidates to be adopted for establishing a de-facto common framework for Big Data intelligent processing. In this sense, the right combination of different Big Data technologies should be selected and adopted, according to their technological needs and specific applications' requirements. Furthermore, it has the potential to organize the way for using geospatial standards across the Big Data processing levels. These concern the *1- pre-processing and exploratory analysis, 2- detection of patterns and objects through advanced mining techniques, 3- Machine learning and (geo)analytics and; 4- Forecasting trends in data with adaptive and controlled uncertainty.* The next challenge is the interpretation of the forecast of trends in Big Data. This requires a good articulation of knowledge expertise while the machine automation of such knowledge is the biggest challenge ahead. Under such circumstances, fundamentals and concepts of AI should be considered.

Challenge-2: Technical priorities

The table below succinctly presents the challenges of EO Big Data, classified into the following categories: Data Management, Data Analytics, Data Advanced Visualization, Data Standardization, Data Processing.

1. Data Management	**2. Data Analytics**	**3. Data Advanced Visualisation**	**4. Data Standardization**	**5. Data Processing**
Scalability	Quality of data	Multiple axis/parameters	Ontologies	Robustness
Adequate and affordable infrastructure to store and process data.	Generating reliable high-quality and affordable information for field management (especially in complex urban systems).	Implementation of 3dimensional visualization services to improve the understanding of spatial relationships between image texture and topography, allowing land use features to be observed not only from the normal vertical view, but also at different scales, and different orientations and perspectives. As example, using this method, the range of sprawl and growth of cities can be identified and determined.	The lack of some standards, particularly to facilitate data analysis at multiscale levels, (e.g. TileMatrixSet).	Data Fusion of multisource data.
Enterprise-class data integration that is dynamic - meeting current and future performance requirements - and extendable - partitioning for fast and easy scalability.	Evidence-based knowledge is required for the management of key public health issues. Detailed travel survey data to capture individual activity-based behavior is required for Human Mobility Patterns Inferred from Mobile Phone Data.	Access to a business intelligence (BI) and analytics data in eye-catching and easy-to-understand formats.	Advanced instruments for territorial planning/management are required. Data set has to be accurate, objective, reliable, comprehensive and always up-to-date, following an integration of multispectral (and hyperspectral) data from various satellites.	Develop automatic or semi-automated processing techniques of very high resolution imagery to successfully provide products that cover extended areas in short time.

Cloud computing platforms to provide virtual clusters and elastic auto-scaling of resources.	The lack of trained people on Innovative Data Analysis (Spatial Analysis and Statistics)	Definition of advanced and integrated data visualization solutions allowing new ways to view EO and non-EO data, through interactive visuals such as bubble charts, word clouds and dynamic geospatial maps.	Implementation of new standards for Big Data (EO, non-EO and fused products) metadata and management.	For multi-temporal monitoring of changes on the land surface implement integrated services with: Image differencing, vegetation index differencing, principal component analysis, direct multi-date unsupervised classification, post-classification change differencing and a combination of image enhancement and post-classification comparison. Develop spectral mixture analysis, artificial neural networks for change detection applications. Developed a curvelet-based change detection algorithm for radar images.
Designing isolation approach to use a data processing design in any hardware configuration without needing to redesign and retune.	Use of multi-temporal imagery approach - combined spring and summer images.	Implementation of customized, innovative, networked and highly affordable visualization solutions (e.g. for Smart City Platforms - indoor and outdoor areas).	Developing interoperable data infrastructure through extensible governance and metadata lifecycle framework.	Developing systems for a stable and fast real-time data handling that can interpret and provide results in real-time.

Table 2 Challenges of EO Big Data within BDVA Reference Model

Challenge-3: Data Analysts and Knowledge Experts

There is clearly a need for knowledge experts to join forces with data analytics experts in order to develop the next generation of Big Data processing systems in the years to come. In order to advance the development of critical decision support which extract timely knowledge form Big Data observation, one would require to tackle the development of the next generation of automated knowledge modelling expert systems. These should encounter the challenges of data ethics, provenance, trustworthiness and uncertainties while processing big data under all levels of the proposed generic framework. Therefore, research programmes which typically address such challenges will be necessary to consider in the near future. This will benefit their multi-disciplinary usage and mapping across multiple domains using EO data combined with air-borne/in situ sensing and social observation data.

The above-mentioned challenges will require increases in the development of hardware and more importantly software which can embed multi-temporal, -spatial and -modal intelligent data analytics, management, re-usable fusion frameworks and supporting knowledge models.

Challenge-4: Data driven services

In order to promote business models that use open geospatial data, public authorities should establish clear guidelines for dealing with privacy issues and comply with applicable data protection laws. In this regard, a challenge is to establish simplified and unified licensing conditions for the commercial use of public sector data. In turn, existing and future acquisition and processing services should ensure aspects such as data availability, updating and quality, as well as the integration of heterogeneous data from multiple sources. In this sense, a challenge is to set global standards for interoperability of data to facilitate their use by these services. On the other hand, the free provision of such data and services requires continued and guaranteed long-term financing. Therefore, innovative public procurement should encourage industry to generate economies of scale in the provision and intermediation of resources and help to optimize investments that allow for the long-term sustainability of these services. Therefore the critical revision of the current approaches and existent Open Data Portals (such as GEOSS) , should be contemplated, to encourage the creation of more effective business services and models based on open data.

Data is always a complicated topic. It is generated, processed and controlled by multiples parties, and generating services that manage different sources of data to produce revenues add an additional layer of complexity.

Three main business models derive from the use of data drive services, data providers (owners of the data that serve it with or without minor modifications), service data providers (they do not need to own the data, they consume it to produce value added services) and consultancy services (based on the knowledge of the data and the user needs act as facilitators).

The main challenges of this kind of services are heterogeneity, lack of structure, error-handling, privacy, timeliness, interoperability and effective dissemination at all stages of the analysis workflow. These challenges are common across the different business models and it is not possible to find a single solution applicable to all of them.

9.2 Conclusions

Examples of directions of interest for Big Data related challenges are: **Satellite Images Time Series, Content and Concept Based Image Retrieval technologies, Data/image Mining techniques and Thematic Exploitation Platforms**.

Therefore, Big EO Data collection, dissemination, processing and added value service delivery for End-Users are an existing technological challenge.

Solutions to the above-mentioned challenges require the combination of space based Remote Sensing with current Big Data driving technologies and open data standards. Besides, an adequate policy framework which facilities its adoption by End-Users in many EU economical critical sectors will be needed. Thus, the benefits from its application will reach out our society.

Synergies of space based Remotes Sensing with Big Data technologies is a need in multiple sectors which requires specific research and continued investment efforts by both private entities and public bodies. Key findings of this white paper are as follows:

- Platforms interoperability needs to be ensured;
- Public End-User adoption needs to be encouraged by market-oriented policies (i.e. governance that promote private downstream services).

As reflected in the document, the current situation provides a very suitable context for the development of a Big Data ecosystem in the field of geospatial data. This type of data presents the four v characteristics of Big Data: volume, variability, velocity and value. In fact, "Big Data is no stranger to the geospatial world; only it is known by several names. Perhaps Big Data is a way of defining a paradigm shift to a data-intensive collaboration where processes reinforce traditional database approaches"[27]. In this sense, the opportunities provided by the European satellite constellation supplying open geospatial data should be exploited by the deployment of the Big Data analytics open services.

There are different technologies developed in other areas that can be applied to the EO data with further developments, between then can be mentioned Machine Learning and in particular Deep Learning, Satellite Image Time Series, Content Based Image Retrieval, Visual Data Mining, 3D Visualization, Streaming in situ technologies or Complex Event Processes.

Most of these mentioned technologies applied to geospatial data require specific infrastructures. There are several initiatives worldwide, both publics and privates, that have already started or they are in preparation for the uptake of EO data into their Big Data infrastructures. In order to have interoperable platforms, the standardization and widespread use of a limited number of these infrastructures would be highly recommended.

The development of a fruitful and efficient ecosystem of Big Data in EO would have a great social impact in such critical areas as Weather and climate forecasts, management of water resources, ensuring healthy and productive oceans, keeping vulnerable ecosystems safe, human health and much more. However, this requires the availability of a large number of data scientists specialized in geospatial data, with solid knowledge in statistical foundations and advanced data analysis methods

[27] https://www.geospatialworld.net/article/big-data-the-future-is-in-analytics

and a thorough understanding of scalable data management, with the associated technical and implementation aspects. To cover this demand the implementation of specifically courses is needed to address the European demand for professionals and researchers with knowledge in big geospatial data exploitation.

Innovation is the driving force of economic growth, and can be looked at from two perspectives. First, innovation in big data driven by industry should be analyzed. Then the perspective of innovations driven by Big Data on industry should be considered. In both perspectives, there are catalysts to be identified and barriers to be named, to allow measures to be taken facilitating the first and eliminating the second.

In the near future, the innovation should focus on EO data feeds in other domains than the 'traditional' ones and on ingesting other available data (than the ones already 'traditionally' integrated) into EO services / applications.

Business opportunities must be promoted to capitalize current EO data; and innovative companies in the EO midstream segment of the value chain need to be strongly supported to ensure reliability, quality and storage of data. Public authorities have a critical role to play in this scenario. Therefore, a long-term vision of Big EO Data policy is essential.

LIST OF CONTRIBUTORS (ALPHABETICAL ORDER)

Name	Marcos Xosé ÁLVAREZ-CID
Position	Technology Transfer Officer
Organisation	FEUGA
Email	mxalvarez@feuga.es

Title and Name	Arne J. BERRE
Position	Chief Scientist
Organization	SINTEF
Email	Arne.J.Berre@sintef.no

Title and Name	Ph.D. Konstantinos CHATZIKOKOLAKIS
Position	Senior Researcher
Organisation	MarineTraffic
Email	konstantinos.chatzikokolakis@marinetraffic.com

Title and Name	Anca Liana COSTEA
Position	Manager, Policies, Strategies and International Cooperation
Organisation	TERRASIGNA
Email	anca.costea@terrasigna.com

Title and Name	Leslie GALE
Position	Marketing & Business Development Manager
Organisation	Space Applications Services
Email	leslie.gale@spaceapplications.com

Title and Name	Ph.D. Angel M. GARCÍA PEDRERO
Position	Postdoctoral researcher
Organisation	Center for Biomedical Technology; Universidad Politécnica de Madrid
Email	angel.garcia@ctb.upm.es

Title and Name	Ph.D. Consuelo GONZALO-MARTÍN
Position	Professor
Organization	Universidad Politécnica de Madrid
E-mail	consuelo.gonzalo@upm.es

Title and Name	PhD. Francisco GUTIERRES
Position	Big Data Analyst
Organisation	Eurecat - Technology Centre of Catalonia
Email	francisco.sacramento@eurecat.org

Title and Name	Codrina Maria ILIE
Position	Senior Scientist
Organisation	TERRASIGNA
Email	Codrina.ilie@terrasigna.com

Title and Name	PhD. Roberta PISCITELLI
Position	Strategy and Policy Officer
Organisation	EGI Foundation
Email	Roberta.piscitelli@egi.eu

Name	Nuria RODRIGUEZ-AUBO
Title / Position	Technology Transfer Officer
Organisation	FEUGA
Email	nrodriguez@feuga.es

Title and Name	Ph.D. Z. A. SABEUR, MInstP, CEng, CPhys
Position	Science Director
Organization	University of Southampton IT Innovation Centre
E-mail	zas@it-innovation.soton.ac.uk

Title and Name	Almudena SÁNCHEZ GONZÁLEZ
Position	Business Development Executive
Organisation	GMV
Email	asanchez@gmv.com

Title and Name	Ph.D. Florin SERBAN
Position	General Manager
Organisation	TERRASIGNA
Email	florin.serban@terrasigna.com

Title and Name	Bernard STEVENOT
Position	Senior Advisor
Organisation	SPACEBEL
Email	Bernard.Stevenot@spacebel.be

Title and Name	Antonio TABASCO
Position	Head of Remote Sensing Applications and Platforms
Organization	GMV
E-mail	atabasco@gmv.com

Title and Name	Ph.D. Dimitris ZISSIS
Position	Head of Research
Organization	MarineTraffic
E-mail	dzissis@marinetraffic.com

ANNEX I – EU DATA POLICY

Copernicus data policy

Regulation (EU) No.911/2010 of 22 September 2010 (2010 GMES Regulation)[28] establishing the GMES (now Copernicus) programme authorized the European Commission to adopt a dedicated data policy. This data policy has been implemented via delegated **Regulation (EU) No.1159/2013 of 12 July 2013** (Copernicus Data Policy)[29]. The Copernicus Data Policy requires Copernicus data and information to be made available on a full, open and free of charge basis, subject to limitations concerning registration, dissemination formats, and access restrictions. The new **Copernicus Regulation (EU) No.377/2014 of 3 April 2014** (2014 Copernicus Regulation)[30], replacing the 2010 Copernicus Regulation, ensures the continuation of this data policy. The key elements of free, full and open access in terms of the Copernicus Data Policy are that 1) there are no restrictions on the use (commercial and non-commercial) nor on users (European and non- European); 2) a free of charge version of any data set is always available on the Copernicus dissemination platform; and 3) data and information are available worldwide without limitation in time. Users shall, however, inform the public of the source of the data and information and notify any modifications made thereto. Limitation of access might be based, on the one hand, on legal grounds such as international agreements, intellectual property rights and the protection of personal data and, on the other hand, on the protection of security interests subject to an assessment of the sensitivity of the data and information to be undertaken by the Commission. Finally, access limitations can be set in exceptional cases for technical reasons, namely if the capacities of the dissemination systems are not sufficient to serve all user requests. Where the Commission restricts access to certain users, those users shall register under a procedure allowing their un- equivocal identification before receiving access.

Sentinel data policy

The European Space Agency (ESA) is developing a family of Earth Observation missions called Sentinels that will meet the operational needs of the Copernicus programme. Each Sentinel mission is based on two satellites to provide reliable data for Copernicus services. Sentinel 1 was launched on 3 April 2014. The Joint Principles for a Sentinel Data Policy were jointly created by the European Space Agency and the European Commission. The Sentinel data policy is set in the context of the Copernicus data policy, and they both require full, free of charge and open data. Sentinel data as well as the products generated from the data and provided via the internet are full, free of charge and open. There is no restriction on who has the right of access to Sentinel data, in particular, no difference is made between public, commercial and scientific use and between European or non-European users. While the data are free of charge there is a registration process and licenses are issued. An exception to open data access is for security reasons.

ESA data policy

The current ESA Data Policy for ERS, Envisat and Earth Explorer Missions (ESA Data Policy) replaced earlier separate data policies and brought the data policy for those Earth Observation missions in line with the Sentinel data policy (see above). It follows the established objective of these missions, i.e. "to maximize the beneficial use of the Earth observation data provided and to maintain a balanced use of these data for a variety of applications, be it scientific, for public good or commercial." ESA EO data from the ERS, Envisat and Earth Explorer missions are available on an open and non-discriminatory basis. The data are split into two major groups: the free data set and the

[28] http://www.d-copernicus.de/sites/default/files/dokumente/gio-verordnung_en.pdf
[29] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R1159&from=EN
[30] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0377&from=EN

restrained data set. Access to the free data set for all types of use is granted upon electronic on-line registration and the user accepting the ESA Terms and Conditions, with immediate on-line access to the data sets thereafter. The restrained data, however, are not systematically processed and made available online. Access to those data is initiated by the submission of a project proposal to ESA, allowing ESA to assign a maximum quota for product delivery. The majority of the data are free of charge, including all the Earth Explorer data. The main element of the restrained data set consists of Synthetic Aperture Radar (SAR) data from the ERS and Envisat missions, plus very large data sets from the satellite missions that require significant ESA resources to compile.

EUMETSAT data policy

EUMETSAT operates in part a full, free of charge and open data policy. The organizations that receive data free of charge are the National Meteorological Services of the EUMETSAT Member States for their Official Duty use. EUMETSAT provides one set of data and services, determined by its Council, on a free of charge and unrestricted basis as "essential" data and products in accordance with the World Meteorological Organization (WMO) Resolution. "Essential" data and products are those required by National Meteorological services to carry out their official duty. Access to these "essential" data and products is granted to all users without a license, without charge and without conditions on use. In addition, certain defined data sets are made available free of charge and no charges are levied for data that are more than 24 h old. For other uses of the data, and especially for commercial uses, fees are charged for EUMETSAT data, products and services, and license conditions are imposed. The EUMETSAT Principles on Data Policy (EUMETSAT data policy) includes a section on data denial in connection with data from the NOAA instruments flown on EUMETSAT satellites where the supply of data can be halted at times of war or other crisis.

SPOT data. In January 2014, the French national space agency (CNES) announced that it would make SPOT satellite data that are more than five years old available to non-commercial users, free of charge. The first tranche of 100,000 SPOT images available under this policy was processed during 2014. Users have to sign a license agreement with SPOT to declare that the use of the images will be for research purposes only. Under the definitions of the Open Knowledge Foundation or the Panton Principles noted in section 1.3 of this paper, the SPOT data are neither open nor full, but SPOT is taking steps in the direction of open data and the data that are provided for research use under the new policy is free of charge.

ANNEX II - EARTH OBSERVATION DATA APPLICATION DOMAINS

The following domains impacted by the use of EO big data were addressed in the present White Paper:

- Agriculture;
- Forestry;
- Safety / Security / Disaster Monitoring / Risk Assessment/ Insurance;
- Energy / Oil and Gas/ Resource Management;
- OCEAN MONITORING: Marine and Maritime / coastal zone management/ Fisheries and Aquaculture;
- Public Health;
- Urban Development/ Smart cities/ Cultural Heritage;
- Transport - navigation.

These domains are briefly described below, from the perspective of the EO data use. The domains description will summarize some generalities, the services and products using EO big data, the typical clients and business models for the EO companies offering the services and products, the needs for cross domain collaboration between the EO experts and the experts from the respective domains, the needs / gaps / ways for raising awareness of the EO Big Data benefits, and related challenges and barriers.

1. AGRICULTURE

Big data in agriculture is a new and not fully explored topic. Using different data and building up different applications for agriculture established the concept of ‘precision agriculture’. Within the last years, considering the opportunities open by big data, the concept moved forward from ‘precision agriculture’ to ‘smart farming’, that goes beyond the primary production and addresses all the food supply chain, from farming operation prediction, real time operational decisions and defining new business models. The farm management needs to address three management functions: monitoring (data gathering), analysis and decision, and intervention.

More and more, the monitoring part of the farm management is done using satellite imagery. This became regularly used for agriculture and is growing in popularity, providing data to enable economic and environmental benefits to farm management processes. Satellite imagery can improve revenue generation for agricultural applications by providing information related to: crops, weather, health status of the plants, draught, fertilizer needs, etc.

For the implementation of the Common Agricultural Policy (CAP), the European Commission requires timely information on the agricultural production to be expected in the current season. This is a central concern of projects like MARS-project (Monitoring Agricultural ResourceS) of the AGRI4CAST and FOODSEC units of the Directorate General Joint Research Center (JRC) of the European Commission.

The JRC’s monitoring of agriculture with remote sensing started in 1988, initially planned to apply emerging space technologies for providing independent and timely information on crop areas and yields. Since 1993, this action has contributed towards a more operative and efficient management of the CAP through the delivery of a broader range of technical support services to DG Agriculture and Member-State Administrations. Since 2000, the expertise in crop yields has been applied outside the EU. Services have been developed to support EU aid and assistance policies and provide building blocks for a European capability for global agricultural monitoring and food security assessment.

1.1 Services / products using EO data

- Crop health monitoring and management;
- Crop type;
- Crop insurance damage assessment;
- Production management practices (e.g. microzoning);
- Fertiliser application requirements;
- Yield estimates;
- Re-growth monitoring;
- Illicit crop monitoring;
- Pest and invasive species monitoring;
- Irrigation requirements and application;
- Field boundary management;
- Field scale mapping;
- Monitoring agri-environmental measures (e.g., acreage) to inform subsidy allocations;
- Assessing storm damage;
- Greenhouse detection.

1.2 Typical clients for EO services and applications

- Farmers / agri-business owners;
- Policy actors – support in European Common Agriculture Policy;
  - National authorities in charge with subsidies payments in agriculture;
  - Eurostat – European Statistics - LUCAS - Land use and land cover survey;
  - Monitoring Agricultural ResourceS (MARS) - European Commission;
- Insurance companies.

1.3 EO companies: business models

- Frame contracts with public authorities;
- Private contracts with single farmers;
- Private contracts with farmer associations;
- XaaS (Anything as a Service): Data as a Service (DaaS), Software as a Service (SaaS), Mapping/Monitoring as a Service (MaaS), etc.

1.4 Collaboration between domain experts and data analysts

- Integration of multisource data: machine generated data (e.g. sensors), social media data, EO data, crowd sensing, etc.;
- Data-driven decision-making processes.

1.5 Interdisciplinary collaboration

- Agronomists;
- Biologists;
- Meteorologists;
- Data Scientists;
- Software engineers;
- Data analysts;

- Economists;
- Geospatial data professionals.

1.6 Awareness / stimulating the use of EO data

- Open geospatial data access;
- Publicly available map servers;
- Platforms and apps for mobile devices that allow crop monitoring, data analytics and remote farm management.

1.7 EO Big data related needs, barriers, challenges

Smart farming implies using big data and therefore big data related challenges arise that had to be tackled. Among these:

- Multisource data integration - heterogeneity of data sources;
- Data access – quick access to relevant data;
- Data analytics: heterogeneity of data semantics; real time processing;
- These techniques are still seen as complex and cost-ineffective by farmers and land owners: need for awareness raising, training and success cases demonstrations.

2. FORESTRY

'Big Data' (geospatial data as well as data from forestry processes and common information e.g. weather data) provides a huge opportunity to increase the efficiency of forest operations. In addition, it adds new possibilities to connect knowledge of basic conditions (e.g. trafficability), efficient forestry and harvesting actions with demand and expectations from forest industries and the society. From the forestry point of view, accurate spatial information makes it possible to move from classic stand-wise management to precision forestry, i.e. micro stand level, grid cell level or tree-by-tree management. Contributing to make key management decisions, such as: what species to plant, when to cut down a specific area or what wood products to cut.

Therefore, the geospatial technologies can improve the current national and European frameworks related to forest management, namely the National Forest Inventory Surveys and the Land use and land cover survey (LUCAS). Both surveys are a rolling programme of forest surveys that collate accurate data on tree size, distribution, composition and condition of forests and woodlands. NFI and LUCAS data is used to assess changes taking place in woodlands over time. These geospatial data and novel big data technologies can be used by researchers and organizations involved in forestry and land management, and in planning, policy development and business.

The challenges of combining forest management and Natura 2000 objectives in forest habitats can be also addressed. It can be done bringing together the latest advances in forest ecology (researchers and scientists) and practice (land owners, foresters, conservationists). Thus, through the integration of geospatial technologies and data, several transversal questions can be addressed, such as:

- What is the latest scientific evidence on the coherence of forest management with Natura 2000 habitats and species conservation?
- How can Natura 2000 conservation objectives be integrated in current forest management and planning?

- What are the European Union legal boundaries for forest management in Natura 2000 sites?
- How the reference values and favorable conservation status of forest habitats should be assessed and monitored?

2.1 Services / products using EO data (+ standards)

- Accurate mapping of forest cover extent;
- Classifying forest species and its quality;
- Determining forest primary productivity;
- Monitoring forest fires;
- Monitoring of deforestation and forest degradation;
- Quantifying forest biomass;
- Performing natural forest inventories;
- Pest and disease control and management;
- Optimize the use resources such as water and energy;
- Monitoring structure and function of Natura 2000 habitats (under the EC Habitats Directive).

2.2 Typical clients for EO services and applications

- Natural resources policy makers;
- Environmental managers and consultants;
- Logging companies;
- Forest Insurance companies;
- Forestry authority expert;
- Forestry landowners;
- Forestry contractor;
- Timber buyers;
- Eurostat – European Statistics - LUCAS - Land use and land cover survey
- European Centre for Nature Conservation (ECNC)
- European Environment Agency (EEA)
- Joint Research Centre (JRC)

2.3 EO companies: business models

- Frame contracts with public authorities;
- Private contracts with private forest owners;
- XaaS (Anything as a Service): Data as a Service (DaaS), Software as a Service (SaaS), Mapping/Monitoring as a Service (MaaS), etc.

2.4 Collaboration between domain experts and data analysts

- Integration of multisource data;
- Generation of prediction models for planning forest work and general management;
- Data-driven decision-making processes;
- Create decision-support tools for efficient, sustainable and value-creating forestry and forest operations.

2.5 Interdisciplinary collaboration

- Forest engineers;
- Biologists;
- Meteorologists;
- Environmental engineers;
- Data Scientists;
- Software engineers;
- Environmental activists;
- Economists;
- Geospatial data professionals.

2.6 Awareness / stimulating the use of EO data

- Open geospatial data access;
- Shared multiuser forest data;
- Publicly available map servers;
- Social forest platforms;
- Providing users tools and new management plans that take into account wood and non-wood products, as well as, conservation areas while at the same time maximizing timber production and economic yield;
- Developing apps for mobile devices that allow field supervision and timber trade.

2.7 Big data related needs, barriers, challenges

Today, foresters are forced to make decisions based either on experience or growth models. In this regard, they need more and better data, and the information extracted from the data, to make better decisions. Industry expects that the use technology and data will create economic benefits similar as the ones reported in the agriculture industry thanks to the precision agriculture.

Unlike the case of the agriculture, the cost of collecting or acquiring forest data is an important issue for the industry and decisions without proper data cannot be relied upon to improve profitability and sustainability.

While technology has been adopted in forest land management, especially geographic information systems and a variety of measurement tools to assess stand growth, the adoption of technology in the registry, especially data analytics for the management and optimization of production systems has room for improvement.

Therefore, one of the major challenges for foresters is to choose sensors wisely and decide which data they should capture given the cost constraint associated with collecting nearly all forms of data. In this regard, the widely use of technologies such as Terrestrial Lidar, Aerial Lidar, Satellite Imaging and Drone based imaging can make much easier for foresters to collect and manage data.

But, collecting only data is not enough, platforms are also needed to help make sense data to enable better decisions. These platforms should allow new data to be easily integrated with existing forest information systems. So that they can take advantage of existing data and also prepare for the massive growth of data coming from field foresters, harvesting machines, remote-controlled aircraft, satellites, and so on.

3. *SAFETY / SECURITY (INCLUDING DISASTER MONITORING / RISK ASSESSMENT/ INSURANCE RISK ASSESSMENT)*

The establishment of Safety and Security is paramount for the protection of citizens, industrial operational activities and critical infrastructure in Europe and beyond. The emergence of high resolution EO data is generating a great resource opportunity for using it to enforce existing means of safety and security monitoring. Nevertheless, the challenge remains for efficiently acquiring such big data and intelligently processing it for real-time critical extraction of knowledge and decision support.

According to the Study on the competitiveness of the EU security industry (ECORYS 2009)[31], the security market is worth 26 billion Euros in Europe with an annual growth rate estimated at 3%. This market covers a wide panel of technologies, where Information Technologies and EO data account for a limited share, along with sensors, screening, communications, protective, intelligence technologies which are dedicated in diverse applications sectors. These include border control, aviation, maritime and terrestrial transports, natural and man-made disasters, utilities, cyberspace and critical infrastrucutre, such as urban infrastructure but partciularly the so called widezones which cross countries aorund the European continent.

Due to the wide scope of the security market, it is necessary, for the purpose of identifying applications, to divide the market into sectors:

Main domains	Sectors
Border control	Border surveillance, terrestrial, marine and maritime, and air space surveillance
Civil Protection	Earthquakes, landslides, floods, submersion, storms, droughts, technological accidents, terrorism, crime
Critical Infrastructure	Utilities, industrial sites, supply chain, transportation systems, widezones
Cyber Security	ICT infrastructure, Data assets and information.
Others	Physical security, protective clothing, etc

Table 3 ***EU Security market size***

The aditional aspect concerns the crisis management cycle, traditionally subdivided into the following elements:

- **Risk Assessment and Prevention**: taking sustained actions for reducing or eliminating long-term risks to people and property from disasters and their effects;

[31] https://ec.europa.eu/home-affairs/sites/homeaffairs/files/e-library/documents/policies/security/pdf/study_on_the_competitiveness_of_the_eu_security_industry_en.pdf

- **Preparedness**: build emergency management functions to respond to any disaster, and recover from it;
- **Early Warning**: provide efficient and reliable information on the occurrence of disastrous event;
- **Emergency response**: conducting emergency operations to save lives, reducing the disaster to acceptable levels, evacuating potential victims, covering the needs of the affected, and restoring critical public services;
- **Recovery**: rebuilding communities in order to return to normal, and protecting against future disaster.

3.1 Services / products using EO data (+ standards)

- Monitoring of Critical Infrastructure (so-called wide-Zones: Motorways; trans-European energy and water pipelines, airports hubs, rail tracks, ports and harbors etc.);
- Surveillance of human behavior (illicit behavior near critical infrastructure);
- Meteorological Forecast (extreme weather forecast);
- Early warning systems (hurricanes, extreme storms, tsunamis, landslides, avalanches, earthquakes, etc.).

Various flagship projects which were funded by the European Commission under the Framework Programme 7 led the basis for product development for the intelligent management of natural crises and extreme events and hazards. These pioneered open service infrastructure solutions for handling ever increasing volumes of observation data, including EO, from heterogeneous sources. These include the following:

TRIDEC[32]: The project focuses on new approaches and technologies for intelligent geo-information management in complex and critical decision-making processes in Earth sciences. The key objective in TRIDEC is to design and implement a collaboration infrastructure of interoperable services through which the intelligent management of information and very large data, dynamically increasing both in terms of size and dimensionality, is efficiently supported. This will enable multiple decision-makers to respond efficiently using a collaborative decision-support environment. TRIDEC will establish rapid and on-demand interoperability of inherited legacy applications and tools owned by the project consortium partners. By using collaborative computing techniques TRIDEC enhances the interoperability of the components to establish a decision-support enterprise system of services which can critically deliver timely information to decision-makers.

ZONeSEC[33]: The project focuses on the deployment of secure and interoperable Big observation data and information management services using open standards. Also, it aims at cost-effectively reusing them in the surveillance of many other European WideZones. These services are part of an advanced Knowledge Base (KB) and primarily focused on large scale surveillance with high performance detection of localized abnormal activities and alerts. Semantically enriched domain knowledge representations shall be stored in the KB for supporting high level data fusion and reasoning with reduced uncertainties and false alerts. Surveillance professionals will securely subscribe to the scalable KB services of the ZONeSEC system of systems with customizable visualization features. Several pilots specializing in the detection of illegal unauthorized entrances to or trespassing premises; or actions to damage to or deployment of harmful devices on installations are demonstrated.

[32] http://www.tridec-online.eu/project
[33] https://www.zonesec.eu/

These concern Water, Oil and Transnational Gas Pipelines; Highways and Rail tracks conveyed in six European countries.

INFRARISK[34]: The methodological core of the project focuses on the establishment of an 'overarching methodology' to evaluate the risks associated with multiple large scale infrastructure networks for various hazards with spatial and temporal correlations. Interdependency is formalized and damage to critical infrastructure is defined in terms of capacity decrements. This is the basis for the development of stress tests for multi-risk scenarios on CIs under extreme hazard rare events. A general framework, is provided with integrated tools for decision making based on the outcome of the stress test. The main objective goal is specifically on the design and development of a strategic INFRARISK Decision Support Tool (IDST) to ensure that the INFRARISK stress tests and the harmonized risks management methodologies and big data analytics modules integrated and driven by a specific process workflow engine. The IDST platform enables risk managers of natural hazards on critical infrastructure access, mine and share data and information, implement multi-risk scenarios and evaluate the cost for full recovery and functional capacities of critical infrastructure.

CRISMA[35]: This project focuses on large scale crisis scenarios with immediate and extended human, societal, structural and economic, often irreversible, consequences and impacts. Typically, these crisis scenarios cannot be managed alone with regular emergency and first responder resources, but require multi-organisational and multi-national cooperation including humanitarian aid. The CRISMA project developed a simulation-based decision support system, for modelling crisis management, improved action and preparedness. The CRISMA System facilitates simulation and modelling of realistic crisis scenarios, possible response actions, and the impacts of crisis depending on both the external factors driving the crisis development and the various actions of the crisis management team.

3.2 Typical clients for EO services and applications

- Homeland security forces (example: The Police force, Homeland Ministry, etc.);
- Civil Protection and Disaster Response Organizations;
- Health care and voluntary care organization;
- Civil Engineering Companies (dredging, bridges maintenance, road management, etc.);
- Operational industries and Agencies (maritime transport, air traffic control, highways agencies, water treatment companies, oil and gas pipelines network operators, etc.).

3.3 EO companies: business models

- Private SMEs to large enterprises processing EO images;
- Public enterprises processing EO Images;
- Collaboration with:
    - National Institutes processing EO images;
    - National Meteorological Office which forecast extreme events;
    - National and Regional Warning Centres.

[34] https://www.infrarisk-fp7.eu/

[35] http://www.crismaproject.eu/

3.4 Collaboration between domain experts and data analysts

- Integration of geospatial data with other sources (social data, native security context information, etc.);
- Generation of human behavior detection and human intentional behavior predictive models (based on observation measurement with context driven knowledge modelling);
- Data-driven decision-making processes based on European Directives on homeland security and the protection of critical infrastructure;
- Acquisition of ETSI (EMTEL) and CEN/CENELEC standards for security regulations and compliance by organizations, especially SMEs.

3.5 Interdisciplinary collaboration

- Computer scientists;
- Chemists;
- Electronics Engineers;
- Civil Engineers;
- Biologists;
- Meteorologists;
- Environmental scientists;
- Data Scientists;
- Psychologists;
- Legal experts (Data Control, processing, ethics and protection);
- Geographers;
- Legal specialists;
- Security ISO standards accreditation experts.

3.6 Awareness / stimulating the use of EO data

There is a need to establish and promote EO under the various regulated ISO standards(ISO 27001-5) for security systems to be taken into account.

One must consider a common generic framework as offering to organisations in the security sector in order for them to adapting EO data streaming into their operational systems for safety and security The emerging big data technologies for EO big data processing and the intelligent extraction of critical knowledge for security and operational decision-support requires new updates on the way the ISO standards for security are compliant and accredited.

3.7 Big data related needs, barriers, challenges

Among the challenges for using EO big data that is encountered in the security sectors, the following are identified below:

- Readiness of EO data catalogue open services for security enterprise;
- Readiness of EO data connectors for security enterprise;
- Cross-border EO data and context information exchange and sharing;
- Systems interoperability across wide zones;
- Scalability of security systems versus EO big data;
- Elasticity of systems and observation sensing networks (EO + in situ observation);
- Data owners - versus - data controllers – versus - data processors;

- Ethics and Data Protection Agreements.

4. ENERGY - OIL AND GAS

The Oil and Gas industry is not an exception when it comes to recognizing the value of big data and the application of big data analytics to understand and improve the performance of aspects of their business and operations. For the O&G this is a very wide span of activities from exploration, discovery to exploitation, operations to decommissioning covering many aspects from the environment to consumer needs. None-the-less it is recognized by the O&G industry that to change the performance on a significant level, analytics needs to be integrated within the day-to-day business operations. In this section, we shall be review how Earth Observations fits into Big Data analytics and critical knowledge extraction that is driving the changes in the oil and gas industry.

The Oil & Gas industry is an active and established user of Earth Observation imagery going back to the early 1970s following the launch of Landsat 1 in 1972.

4.1 Services / products using EO data (+ standards)

There are plenty of sources of information that can be accessed to learn about how and where EO can be applied. One recent major study is the in 2014 awarded contract from the European Space Agency (ESA) and the International Association of Oil and Gas Producers (IOGP) that funded 4 projects investigating the needs and solutions of the oil and gas industry for geospatial products coming from satellite data. The 4 projects; 2 focused on on-shore and 2 focused on off-shore were run together under the name of EO4OG (Earth Observation for Oil and Gas).

The projects identified 15 base geo-information needs, see *Table 4*.

Base geo-information requirement	Description
Air quality and emissions	NOx, SOx, CO_2, particulates emissions and concentration. Methane emissions.
Distribution and status of assets	People, equipment, buildings.
Distribution and status of infrastructure	Pipelines, roads, and railways and their condition.
Distribution of habitat and biodiversity	Critical habitat for important flora and fauna. Distribution of important flora and fauna.
Land cover	General land cover information and specific thematic classes such as forest types and parameters.
Land use	General land use information and specific thematic classes such as urban and agricultural land use, as well as other industrial development.
Lithology, structural geology, surficial geology	Surficial geology, geo-hazards
Base images	Multi-spectral images that typically have precise ortho-correction to support multiple uses.
Sub-surface features	Evidence of sub-surface infrastructure.
Surface motion (horizontal and vertical)	Land surface deformation related to seismic activity or production, as well as environmental change such as permafrost degradation.
Terrain information	Stability, surface properties and characteristics
Topographic information	Elevation, slope, and other derived measures.
UXO hazard	Estimates of UXO contamination.
Water quality	Water quality information for rivers and lakes, such as nutrients, sediment concentration, temperature, metals, etc.
Water quantity	Water distribution, including extent of lakes, streams, and wet areas, flood extent, snow and ice cover extent and conditions.

Table 4 Base geo-information needs from (Hatfield Consultants Partnership In Association With: Arup, Rps Energy, C-Core, Centrum Badań Kosmicznych (Poland), July 2014)

The operational challenges are organized in 5 themes:

1. Land Seismic Survey;
2. Surface Geology Characterization;
3. Land Surface Deformation;
4. Environmental Risks Management;
5. Logistics, Planning and Operations.

In addition, we consider that the operational challenges should be extended to include:

6. Platforms decommissioning;
7. Deep sea exploration and mining.

The base geo-requirements have been mapped to the themes as shown in *Table 5*.

	Air quality and emissions	Distribution and status of assets	Distribution and status of infrastructure	Distribution of habitat and biodiversity	Land cover	Land use	Lithology, structural geology, surficial geology	Base images	Sub-surface features	Surface motion (horizontal and vertical)	Terrain information	Topographic information	UXO hazard	Water quality	Water quantity
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Seismic Planning/Survey	▲	▲	●	●	●	●	▲	●	●	▲	●	●	●	▲	●
Surface Geology Characterization					●	●	●	●				●		●	●
Land Surface Deformation	▲	▲	▲	▲	●	●	▲	▲	▲	●	●	▲	▲	▲	●
Environmental Management	●		●	●	●	●		●		●	●	●		●	●
Logistics, Planning and Operations	▲	●	●	▲	●	●	●	●	▲	●	▲	●	▲	▲	●
Platform decommissioning	●	●	●	●	●	●	●	●			●	●		●	●
Deep sea exploration and mining	▲	●	●	●	▲	▲	▲	●	●	●	●	▲	▲	●	●

Table 5 : Mapping Base Geo-information Requirement to Themes from (Hatfield Consultants Partnership In Association With: Arup, Rps Energy, C-Core, Centrum Badań Kosmicznych (Poland), July 2014) [

This was further evaluated leading to a total of 224 challenges faced by the O&G industry where there is a potential for satellite EO data to play a role in providing a solution. From the challenges, 94 products were defined.

The base geo-information needs follow the phasing of an O&G project however the demand for information in terms of data volume, resolution, precision, frequency, and coverage changes can vary depending upon location, environmental constraints and operational risks.

Today, Copernicus with its Sentinel missions and new commercial satellites offer an unprecedented wealth of information. This is particularly relevant for the operational needs, however the importance and value of baseline data and information (hindcast data) should not be forgotten since historical and baseline information is required prior to oil and gas activities. Baseline data is often used to analyze the impacts of oil and gas activities, as well as to manage potential external effects on infrastructure and assets. With the Sentinel missions assured the discontinuity of data arising from satellites with limited life has been eliminated.

Consideration is needed for standards defining how geo-information products are developed, their content, applicability and quality to make it easier to compare between products and select the most appropriate product or service based on the business need.

Lastly, there is a perception that there is a degree of over-complication in the specification of the products and services for obtaining the data leaving customers confused by the number of available parameters and configurations.

All the information coming from the extensive analysis performed in EO4OG can be found on the portal https://earsc-portal.eu/display/EO4/About+EO4OG , developed by EARSC following the results of the projects and presenting the information in a structured way across all 4 projects.

4.2 Typical clients for EO services and applications

Besides the O&G major companies, the O&G industry has an extensive set of potential clients for EO services and applications that can be addressed that are directly involved in the activities of the O&G major companies (1, 2 and 3 Tier companies).

In addition, there are related companies and organizations, particularly in the area of monitoring, policy and health, safety and environment.

4.3 Business models

The O&G industry has a well-organized procurement process that foresees:

- Direct hire of contracted consultancy;
- Product purchasing;
- Delivery on consultancy project basis;
- Multi-year service contracts;
- Exploratory drilling governments agencies licensing;
- Operations and production government agencies licensing;
- Decommissioning of production platforms government agencies licensing.

4.4 Collaboration between domain experts and data analysts

- Multisource data provisioning for integrated process workflows;
- Generation of prediction models for planning and operations (logistics);
- Data-driven decision-making processes;
- Environmental monitoring;
- Crisis management;
- Surveying (Deep sea exploration, National Geological Surveys…etc.);

4.5 Interdisciplinary collaboration

- O&G engineers;
- Chemists;
- Biologists;
- Meteorologists;
- Environmental engineers;

- Environmental consultants;
- Health and safety executives;
- Data Scientists;
- Software engineers;
- Environmental activists;
- Economists;
- Geospatial data professionals;
- Legal specialists;
- Governments' representatives;
- Environmental regulators (EEA, IMO, OSPAR etc..).

4.6 Awareness / stimulating the use of EO data

It is interesting to observe that not one major O&G Company is a member of BDVA.

For O&G we need to develop and promote EO in a new framework that takes into account:

- The changing business model of O&G companies where full vertical integrated companies are making way for a partnering and collaborative working;
- The operational models that require quick to develop and short times to operation for services.

We must be aware that organizations have widely varying approaches to using EO data and integrating the data into their processes;

- O&G majors with their own teams of experts and are able to undertake research and development and product development;
- O&G that have their own capabilities but make more use of contracted services;
- Smaller companies buy when needed from EO service providers.

To increase the visibility of EO possibilities to the O&G industry we need to align more strongly with organisations such as EARSC who represents a large number of service companies, many of which are involved in O&G. There is also the need to investigate the possibilities of Joint Industrial Projects alongside EC and other initiatives. For example, there are several decommissioning programmes which are planned to go ahead in the coming decade (up to 2025) require environmental monitoring and assessment of the environmental risks which may be entailed during decommissioning. These can be enhanced and supported with the use of EO and big data technologies. For the European context one may principally focus on the ongoing decadal North Sea offshore platform decommissioning projects. (See https://oilandgasuk.co.uk/decommissioninginsight.cfm)

4.7 Big data related needs, barriers, challenges

We can expect that as Big Data is becoming more critical to the decision-making process for O&G companies the need for data will grow and the level of integration requirement will increase.

However, concerns are voiced that the procedures and processes for using EO data are still too challenging. In order to increase and further develop the added value services based on EO the data needs to be prepared, connected and ready for use with the tools and infrastructure that the O&G organizations use for business analysis and operations. This is sometimes referred to as `Analysis

Ready Data'. Better knowledge is required to stimulate the use of EO data in the "Big Data" processes of the O&G industry.

European initiatives are focused on making data accessible. For the O&G companies this is not sufficient. There is a need for a better awareness of the O&G data needs and formats they require for integration into their GIS and other information systems. The IOGP (International Association of Oil & Gas Producers, http://www.iogp.org) whose members produce 40% of the world oil & gas would be a good starting point to discover how we can better serve the O&G industry.

5. MARINE AND MARITIME (INCLUDING COASTAL ZONE MANAGEMENT/ FISHERIES AND AQUACULTURE)

Climate change, as a known scientific fact which is highly supported by rigorous long term observations and monitoring, has been exacerbated by human activities in the last two centuries or so. The industrialization of many processes for large scale food production, transports and trades around the world are all associated with the extensive usage of our marine and maritime sectors as means of resourcing and pathways for transferring goods and commodities around the globe. This has led to high levels of pollution across our oceans with evidenced damages of natural marine assets and resources. This has put very important marine ecosystems into dangerous levels of potential destruction to total extinction. The majority of the population in Europe heavily relies on marine and coastal resources. Nevertheless, one stresses that in order to make Europe's marine and coastal resources sustainable, it will be important to achieve the following:

1- Monitor and raise standards of water quality of the coastal and marine environments
2- Reduce the risks of industrial pollution
3- Designate more protected zones for marine species and ecosystems
4- Regulate the fishing industries' operations for assured sustainability with more rigorous rules that are permanently supported by continuing integrated scientific studies, conducted at regional, national and European levels
5- Preparedness of coastal authorities for rapid response to marine and maritime disasters
6- Advances in engineering and technologies for the development of smarter 21st century coastal defenses to sustain the effect of climate change (rapidly rising sea levels, more frequent extreme events, storms and flooding)

5.1 Services / products using EO data (+ standards)

- Monitoring of coastal water quality and compliance with the European Water Framework Directives (and bathing Water Quality Directive)
- Meteorological Forecast (extreme weather forecast, Flood alerts)
- Algae Bloom observation/detection and risk alerts
- Monitoring of coastal erosion and sediment transport trends
- Monitoring marine habitats of endangered marine species (example: sea Mammals, bluefin tuna, sea turtles etc.)
- Monitoring of ships atmospheric pollution
- Ship Voyage and Routing Management
- Ship emission and Fuel Consumption Optimisation (SOx, NOx etc.)
- Ship Traffic Management (Safety and Risks of Collision)
- Port and Harbours Operations

5.2 Typical clients for EO services and applications

- Port Authorities (ship navigation, shipping routes…)
- International Maritime Organisation
- Lloyds Register
- Fishing industries (protected zones, fish stocks management…)
- Aquaculture industries (microbial risks contaminations, water quality etc.)
- Tourist Industries (wáter quality, bathing water quality forecast…)
- Maritime Coast-Guard Agencies (Search and Rescue, clean-up operations….)
- Shipping Industries (Voyage optimisation, fuel consumption management…)
- Environment Agencies (water quality management and protection of marine fauna and flora)
- Marine species Watch (seabirds, Sea mammals sanctuaries management…etc.)
- Ship Managers
- Charterers
- Cargo Consignees
- Insurance companies

5.3 EO companies: business models

The Integrated Applications Promotion (IAP), ARTES Satcom Applications (SATCOM-APPS) and Automatic Identification System (SAT-AIS), under ESA and H2020 programmes, are all engaging with several national and international end-user organisations. These expressed great interest in satellite based services to complement and enhance their current capabilities that are relevant to their daily operations and tasks. Furthermore, they are composed of the following:

- Private SMEs to large Entreprise processing EO images
- Public Entreprise processing EO Images
- National Institutes processing EO images
- National Meteorological Offices for the forecast of extreme events
- Others

In context of the above, many research and development projects focusing on the use of EO / Remote Sensing and big data technologies. These can be identified as flagship projects for new business models concerning the use of EO in the space domain for multiple applications. Among these projects, one includes:

- WONDER[36] - The project aims at developing an automated, cloud-based digital service for environmental monitoring which could be utilized globally. The solution builds on earth observation, global positioning, crowd sourcing, internet of things (IoT) and predictive analysis.
- SYMPA[37] - can be defined as an “All-around solution for the sustainable monitoring and control of Marine protected Areas”, Vitrociset Belgium, le Centre Spatial de Liège, and LaMMA are in fact developing a system for the supervision of the touristic and commercial traffic going through Marine Protected Area. Sympa service offer is completed by adding a “Water quality” monitoring which integrates EO data into innovative biogeochemical and hydrodynamic models as support of a sustainable and competitive tourism.
- EO CROWD[38] - Over 90% of our seas are unsurveyed. EO Crowd uses crowd sourcing, EO data and open data to fill this data gap.

[36] https://business.esa.int/projects/wonder
[37] See: https://business.esa.int/projects/sympa
[38] https://business.esa.int/projects/eo-crowd

- ICWM FOR MED[39] - ICWM for MED solution provides water quality monitoring and surveillance of coastal areas by integrating information from: Earth Observation, Navigation and Communication satellite assets; in situ measurements; crowdsourcing information. It intends to address the current limitations from both Earth Observation based and traditional at sea measurements methods in order to support the duties of environmental agencies and coastguard.
- BLUE BELT[40] - Embedded into the SAT-AIS context, this IAP Demonstration Project aims to support the EMSA Blue Belt project. Providing space-based AIS data, it seeks to complement the terrestrial AIS data and provide an added-value source of information for Customs Authorities interested in maritime situational awareness for vessels and geographical focus areas.
- EO4wildlife[41] focuses on the intelligent management and processing of EO big data together with the implementation of new advanced data analytics and Knowledge Base services. These specialise in the prediction of new marine wildlife migratory behaviour which may be due to the changing ocean resources and climate conditions around the world. The research is currently leading on the implementation of web enabled advanced data analytics open services which comply with OGC standards. Big data connectors and a catalogue service is being installed to allow access to the EO big data heterogeneous sources from COPERNICUS satellite.
- SEDNA[42] ("Safe maritime operations under extreme conditions: the Arctic case") is a H2020 research project that is developing an innovative and integrated risk-based approach to safe Arctic navigation, ship design and operation using a big data knowledge and analytics services. These use EO and in situ data including metocean data, sea ice, shipping routes and IMO regulations for safety in the arctic regions.

5.4 Collaboration between domain experts and data analysts

- Oceanographers
- Marine scientists
- Naval Architects
- Marine Engineers
- Data scientists

5.5 Interdisciplinary collaboration

- Computer scientists
- Marine Chemists
- Sensors Electronics Engineers
- Coastal Engineers
- Marine Biologists
- Oceanographers
- Naval Architects
- Meteorologists
- Environmental scientists

[39] https://business.esa.int/projects/icwm-for-med
[40] https://business.esa.int/projects/blue-belt
[41] www.eo4wildlife.eu

[42] https://sedna-project.eu

- Data Scientists
- Maritime Archeologists
- Legal experts (Maritime heritage, Maritime Law, IMO)
- Geographers

5.6 Awareness / stimulating the use of EO data

The most recent advances in smart sensing and big data technologies are making Remote Sensing systems more attractive for assessment of marine (and maritime) information systems. This is specifically for monitoring the oceans, open seas, wetlands, estuaries, and coastal environments. These water environments exhibit their own respective high spatial complexities and temporal dynamic processes. Following the eminence to act on the effects of climate change, it is highly important to monitor such important water environment for assuring their protection and sustainability of usage. Their importance as socio-economic assets for Europe and beyond requires their comprehensive protection using advanced EO resources.

One important element here for stimulating the use of EO data for the marine and maritime sector is really about bringing it nearer to existing marine observation systems across Europe while maximising its usage across the board by the communities concerned. These are mostly using monitoring systems which are based on first principle oceanography based models that have no adaptive functionalities to some extent for the accurate forecast of large scale integrated marine environment processes. Others rely on human volunteers for the monitoring of species, which is not an adequate way to monitoring marine life either. Furthermore, ship navigation systems require more accuracy in shipping routes as well as their environmental attributes in real-time in order to improve their management of ship fuel consumption, voyage optimisations and safety. The presence of sea-ice and the identification of safer routes in extreme marine environments (i.e. Arctic shipping) and the prevailing weather conditions are crucial parameters for safer navigation. Shipping routes for safer transport and efficient use of fuel are critical for worldwide commerce. EO supplies of critical information is the way forward for improving maritime transport in the new emerging Arctic routes. These are becoming more of a possible reality due to climate change, while they have a potential of reducing ship emission along the Asia-Europe-America transport.

As an example, Finland is 90% dependent on sea transport for its exports and imports. In this context, reliable and timely monitoring of icebergs and sea's ice thickness to support icebreakers is key for the economy of the country. A recent study[43] from EARSC and ESA estimates that between EUR 24 million and EUR 116 million are generated in Finland and Sweden thanks to the use of satellite radar imagery. Having access to an archive of old data does not bring any value for icebreaker activities. Disaster monitoring and response is another example in which reliable data acquired timely are critical for safe transport. Besides the Baltic Sea, maritime activities through the 'Northern Sea Route' above Russia have also increased due to climate change. It is noteworthy that 2017 was the first year that a tanker managed to cross this route without the assistance of ice breaker. The economic impact of this route is quite important, as it cuts transit times for ships travelling between Pacific and Europe by around ten days and provides access to Arctic oil fields. Reliable and timely EOs will boost significantly the shipping activities of this route providing safe routing options for tankers, cargos and any other type of merchant ships.

[43] http://earsc.org/news/between-24m-and-116m-per-annum-of-economic-value-generated-thanks-to-the-use-of-satellite-radar-images

All these requirements by the concerned industries can be met by the use of EO data and aggregating then with existing in situ data for advancing forecasting models with much greater controlled uncertainties. These will give better confidences for operational decision-support as a result.

5.7 Big data related needs, barriers, challenges

Among the challenges for using EO big data that is encountered include the following:

- Readiness of EO data catalogue open services for marine and maritime industries
- Readiness of EO data connectors for marine and maritime industries
- Systems interoperability across maritime borders
- Scalability of systems versus EO big data
- Elasticity of systems and observation sensing networks (EO + in situ observation)
- Data owners - versus - Data Controllers – versus - Data Processors
- Maritime law, IMO regulations knowledge base and models for advancing marine decision-support systems.

6. PUBLIC HEALTH

One emerging domain where EO Big Data can contribute is the public health. More and more services making use of the information that can be extracted from this data become of relevance for issues like prevention, early detection of infectious diseases or for diseases monitoring.

6.1 Services / products using EO data (+ standards)

- Monitor air pollutants;
- Harmful algal blooms;
- Monitoring of disease vectors;
- Early detection of infectious diseases;
- Diseases surveillance;
- Early detection of droughts and famines.

6.2 Typical clients for EO services and applications

- World Health Organization;
- National and regional health systems;
- Environmental managers;
- Supranational, national and regional policy makers.

6.3 EO companies: business models

- Private companies to give service to public bodies;
- Public agencies (health, environment, etc.)

6.4 Collaboration between domain experts and data analysts

- Integration of geospatial data with other sources (social data, health individual data, clinical data, etc.);
- Generation of predictive models (particle entrainment and atmospheric pollution, harmful algal blooms, etc.);

- Data-driven decision-making processes in the health sector;
- Development of visual analytics tools.

6.5 Interdisciplinary collaboration

- Physicians;
- Epidemiologists;
- Environmental engineers;
- Data Scientists;
- Software engineers;
- Geospatial data professionals.

6.6 Awareness / stimulating the use of EO data

- Open geospatial data access;
- Developing apps for mobile devices that implied citizens in data collection and provide them useful health information;

6.7 EO Big data related needs, barriers, challenges

To realize the full benefits of using remote sensing data for public health, an unlimited and lower cost access to the high resolution satellite remote sensing data is necessary.

Data standardization. The output of spatial data integration and analysis will need to be translated into formats useful for a wide range of public health users and accessible to multiple agencies and jurisdictions.

- It is necessary to provide priority and support **for consortia of producer and public health agencies** to pursue the issuance of collaborative solicitations to further mature application activities, initiate new ones, and make mature applications operational. This facilitates the integration of Earth observations from space, *in situ*, epidemiological and social data to relate exposures with health outcomes.
- The private sector should be engaged as an active participant in the application of Earth observations to public health. Innovative public-private partnerships at the different levels should be promoted.

7. URBAN DEVELOPMENT (INCLUDING SMART CITIES AND CULTURAL HERITAGE)

Urbanization is not merely a modern process, but a rapid transformation of human social roots on a global scale, where village culture is being rapidly replaced by the urban culture. Therefore, it is necessary and fundamental for policy makers to incorporate big spatial data technologies into urban planning and management. Urban planning involves a large volume of geospatial data both at the time of planning and at the time of implementation of the plan to determine the status of the available facilities. Thus, remote sensing techniques provide accurate, orderly and reliable information for planning and management of a city. For instance, Remote Sensing is extremely useful for change detection analysis and selection of sites for specific facilities such as hospitals, cultural heritage, restaurants, solid waste disposal and industry.

Additionally, the current knowledge base at the intersection of cultural heritage and smart cities can be enriched. Smart cities stand for a conceptual technology-and-innovation driven urban development model. Historical and cultural heritage of cities is and can be underpinned by means of smart city tools, solutions and applications. Therefore, the cultural heritage management can be incorporated in several different strategic areas of the smart city, such as Barcelona, Amsterdam, and London, reflecting diverse lines of thinking and serving a range of aims, depending on the case. Although potential applications and approaches abound, cultural heritage presently stands for a mostly unexploited asset, presenting multiple integration opportunities within smart city contexts. In fact, the cultural heritage is not systematically exploited and formally incorporated in smart city initiatives, despite the fact that it offers an array of opportunities for smart city development, while important technological developments have taken place in the cultural heritage domain in recent years.

7.1 Services / products using EO data (+ standards)

- Accurate mapping of Infrastructure facilities in urban areas;
- Urban sprawl analysis and urban land use change analysis;
- Simulation and forecasting (indoor and outdoor Urban Heat Island (UHI));
- Urban Tree health inspection;
- Urban canopy cover;
- Tree inventory creation and update;
- Irrigation pattern classification;
- Environmental Impact; climate variables (C02, Temperature, etc.);
- Smart Urban management;
- Smart mobility applications/journey planner;
- Interactive tourist attractions maps;
- Citizen-driven innovation (mainly participatory sensing for environmental monitoring);
- Integrated wayfinding navigation system (journey planner) including points of interest.

7.2 Typical clients for EO services and applications

- Metropolitan Planning Agencies;
- Urban planning authorities and agencies - municipalities;
- Services and public benefits urban dwellers;
- Households (census) managers – National Statistics Institutes (NSI);
- Air quality managers;
- Health policy makers;
- Technological centers and private companies – micro-mobility patterns in cultural heritages (e.g. Gaudí´s emblematic temple “Sagrada Familia” in Barcelona) and commercial surfaces (e.g. Lidl);
- Research Institutes, Centres and Groups – Universities.

7.3 EO companies: business models

- Private innovative companies to give new solutions that can be used by decision makers in cities;
- Urban planning authorities and agencies are improving the quality of infrastructures (buildings and cultural heritage), energy efficient services provision and environmental conditions.

7.4 Collaboration between domain experts and dat analysts

- Integration of geospatial data with other sources (Household(census) data, Social data, health individual data, house renting data, etc.);
- Generation of predictive models (C02 uptake and Heat reduction potential, etc.);
- Data-driven decision-making processes in the urban sector;

- Development of visual analytics tools (2D and 3D) – Responsive and Augmented Geovisualization and Augmented reality data visualization application for smartglasses (Microsoft Hololens);
- Multisource data provisioning for integrated process workflows;
- Generation of prediction models for planning and operations (urban facilities and enterprises logistics);
- Environmental monitoring (urban forest, urban ecosystems, air quality, etc.);
- Improvement of Indoor positioning and micro-mobility tools;
- Risk Assessment (Managing Terrorism Risk).

7.5 Interdisciplinary collaboration

- City planners;
- Architects;
- Urban activists;
- Environmental engineers;
- Data Scientists;
- Software engineers;
- Telecommunications engineers;
- Geospatial data professionals.

7.6 Awareness / stimulating the use of EO data

- Open geospatial data Access;
- Promote a Collaborative Urban Development;
- Developing apps for mobile devices that implied citizens in data collection and provide them useful urban information.

7.7 EO Big data related needs, barriers, challenges

A fundamental challenge with use of urban remote sensing and big data technologies in understanding urban environments is the need to overcome barriers associated with the inability of the potential users of the big data technology and products derived from it. Traditionally these barriers have included cost, lack of adequate computer hardware and software, access to the data and satellite imagery, and limited expertise and knowledge to understand and make use of the resulting data and related products, such as the Urban sprawl, Urban Heat Island (UHI) and urban land use change mapping.

There is a need for more European and local companies to supply urban monitoring products through friendly user interfaces, after which the availability of new data and consequently new products can be expected to lead to an increase in their professional and research activities.

ANNEX III - COPERNICUS CORE SERVICES

Copernicus Atmosphere Monitoring Service (CAMS[44])

CAMS is operational from July 2015 and it provides data and information on atmospheric composition, describing the current situation, offering forecasts and making analysis on retrospective data records for recent years. The service monitors and forecasts parameters like: greenhouse gases, reactive gases, ozone and aerosols. It provides analysis and forecasts for the European air quality, and offers information on the solar radiation resources at the Earth's surface. In addition, CAMS compiles emissions inventories that support the modeling and estimation of the CO2 and CH4 fluxes at Earth's surface. The domains that benefit from using CAMS via different applications range from environmental monitoring, renewable energies, meteorology and climatology to health.

Copernicus Marine Environment Monitoring Service (CMEMS[45])

Operational since May 2015, the service focus on the status and dynamics of the ocean and marine ecosystems for the European seas and the global ocean. CMEMS delivers data and information (including forecasts) that support all marine area applications: from maritime operations (transport – ship routing, search and rescue, marine safety), marine resources (aquaculture, fishery), coastal and marine environment (water quality monitoring, pollution control, coastal erosion, sea temperature monitoring). The service provides also high relevant data and information for weather, climate and seasonal forecasting (temperature, salinity, currents, wind, sea ice).

Copernicus Land Monitoring Service CLMS[46]

The service is operational since 2012 and it offers information on land cover and related parameters that gives information like the vegetation state or the water cycle. CLMS comprises three components:

- A global component – providing biophysical parameters worldwide, describing the state of vegetation, the energy budget (e.g. albedo, land surface temperature, top of canopy reflectance) and the water cylcle;
- A Pan-European component – it will produce high resolution data sets describing the main land cover: artificial surfaces, forest, agriculture areas, wetlands and small water bodies;
- A local component – it provides specific and more detailed information complementary to the Pan-European component, focused on identified hotspots (big European cities, riparian areas, etc) prone to different environmental challenges.

Copernicus Climate Change Service (C3S[47])

C3S is under implementation (the beta version of the service portal is available), and is based on the outcomes of a series of projects funded under the 2013 FP7 Space Call related to climate modelling and observation analyses. The service addresses the environmental and societal challenges related to the climate changes associated with human activities. The service will provide climate indicators

[44] http://copernicus.eu/main/atmosphere-monitoring
[45] http://copernicus.eu/main/marine-monitoring
[46] http://copernicus.eu/main/land-monitoring
[47] http://copernicus.eu/main/climate-change

and climate indices, supporting monitoring and prediction of the climate changes in the view of mitigation and adaptation.

The service will provide access to several climate indicators (e.g. temperature increase, sea level rise, ice sheet melting, warming up of the ocean) and climate indices (e.g. based on records of temperature, precipitation, drought event) for both the identified climate drivers and the expected climate impacts.

Copernicus emergency management service (EMS[48])

EMS comes in support to the management of natural disasters, emergency situations and humanitarian crisis, with timely and reliable geo-spatial information from satellites, in situ data and from open data sources. The service comprises a mapping component, with worldwide coverage and an early warning component.

The mapping component is operational since 2012, and the products it delivers are maps based on satellite images, that can be used as such or integrated with other data sources to support the decision making process in all the emergency management cycle (preparedness, prevention, disaster risk reduction, emergency response and recovery).

The early warning component includes two systems: the European Flood Awareness System (EFAS), providing information and forecasts related to floods and The European Forest Fire Information System (EFFIS), giving information on forest fires and forest fire regimes in the European, Middle Eastern and North African regions.

Copernicus service for Security applications[49]

The services addresses Europe's security challenges, providing information in support of crisis prevention, preparedness and response improvement in three areas:

- border surveillance – the focus is on the reduction of the number of illegal immigrants entering EU undetected, increasing the number of immigrants rescues on the sea and prevention of cross-border crime;
- maritime surveillance – the service comes in support of maritime security objectives and related activities in the maritime domain, referring to navigation, fisheries control, marine pollution, and law enforcement;
- support to EU External Action – supporting EU in assisting third countries in crisis situation and to prevent global and trans-regional threats having a destabilising effect

[48] http://copernicus.eu/main/emergency-management
[49] http://copernicus.eu/main/security

ANNEX IV - ESA THEMATIC EXPLOITATION PLATFORMS

Geohazards TEP[50] – GEP - aims to support the exploitation of satellite EO for geohazards. It represents a continuation of the Supersites Exploitation Platform (SSEP), originally initiated in the context of the Geohazard Supersites & Natural Laboratories initiative (GSNL). GEP as well as SSEP platform allow access to data, tools and processing including INSAR.

Coastal TEP[51] - C-TEP is a data access service dedicated to improving the efficiently of data-intensive research into our dynamic coastal areas. Its objective is to provide a virtual platform where EO data and in-situ coastal data may be accessed, processed, analyzed and shared for the enhancement of our maritime management, research and R & D sectors. C-TEP will provide a virtual platform for users interested in deriving information from EO data for a wide range of coastal fields of study e.g. climate change, economic development, habitat classification and coastal research. It does this by facilitating access to a range of key data and cutting edge processing tools through a cloud-based workspace.

Forestry TEP[52] - The F-TEP vision is to be a one-stop shop for forestry remote sensing services for the academic and commercial sectors. The service offers access to pre-processed satellite and ancillary data, computing power, and software access and hosting. The Platform will also offer access to commercial software and services. The aim is to evolve the Platform offering services provided by the European value adding SME's.

Hydrology TEP[53] - A portal providing LARGE SCALE EO SERVICES & PRODUCTS customized for hydrology applications. Flood monitoring and small Water bodies mapping, Water quality and level, Hydrological models.

Thematic Apps:

- Hydrological Modelling;
- Water Level;
- Water Quality;
- Flood Monitoring;
- WOIS – Water Observation and Information System.

Polar TEP[54] - will provide polar researchers with access to computing resources, earth observation (EO) and other data, and software tools in the cloud.

Urban TEP - main goal is the implementation of an instrument that helps addressing key research questions and societal challenges arising from the phenomenon of global urbanization; allows users to effectively utilize Earth Observation (EO) imagery and existing auxiliary data (e.g., geo-data, statistics) to measure and assess key properties of the urban environment and monitor the past and future spatiotemporal development of settlements.

Thematic Apps:

[50] https://geohazards-tep.eo.esa.int/#!
[51] https://coastal-tep.eo.esa.int/portal/
[52] https://forestry-tep.eo.esa.int/
[53] https://hydrology-tep.eo.esa.int/
[54] p-tep.polarview.org/

- Global Urban Footprint 2012 – Vector;
- Global Urban Footprint 2012;
- TimeScan Landsat 2015;
- Data Subsettig Service;
- Worldpop.

ANNEX V – EC AND ESA EO BIG DATA PROJECTS

1. Ongoing and completed EO projects – European Commission

Call	Topics	Budget	Funded projects
H2020-EO-2014-2015 (closed)	EO-1-2014: New ideas for Earth-relevant space applications (RIA)	46,5M€	SPICES BACI URBANFLUXES EUSTACE EGSIEM
	EO-2-2014: Climate Change relevant space-based Data reprocessing and calibration (RIA)		FIDUCEO
	EO-3-2014: Observation capacity mapping in the context of Atmospheric and Climate change monitoring (IA)		GAIA-CLIM
	EO-1-2015: Bringing EO applications to the market (RIA)		APOLLO BASE-platform EOMonDis EONav
	EO-2-2015: Stimulating wider research use of Copernicus Sentinel Data (RIA)		Co-ReSyF EO4wildlife AURORA MULTIPLY
	EO-3-2015: Technology developments for competitive imaging from space (RIA)		S3NET ONION
H2020-EO-2016-2017 (closed)	EO-1-2016 and EO-1-2017: Downstream applications (IA)	45M€	SPACE-O EUGENIUS DIANA CyanoLakes
	EO-2-2016: Downstream applications for public sector users (PCP)		MARINE-EO
	EO-3-2016: Evolution of Copernicus services (RIA)		E2mC SENSAGRI CEASELESS ECoLaSS Copernicus App Lab
H2020-EO-2018-2020 (open)	**DT-SPACE-01-EO-2018-2020: Copernicus market uptake**	87M€	
	DT-SPACE-07-BIZ-2018: Space hubs for Copernicus		
	LC-SPACE-02-EO-2018: Copernicus evolution – Mission exploitation concepts		
	LC-SPACE-03-EO-2018: Copernicus evolution - Preparing for the next generation of Copernicus Marine Service ocean models		
	LC-SPACE-14-TEC-2018-2019: Earth observation technologies		

Call	Topics	Budget	Funded projects
H2020-EO-2018-2020 (forthcoming, 2019)	DT-SPACE-01-EO-2018-2020: Copernicus market uptake		
	DT-SPACE-06-EO-2019: International Cooperation Copernicus – Designing EO downstream applications with international partners		
	DT-SPACE-09-BIZ-2019: Space hubs (support to start-ups)		
	LC-SPACE-04-EO-2019-2020: Copernicus evolution – Research activities in support of cross-cutting applications between Copernicus services		
	LC-SPACE-05-EO-2019: Copernicus evolution – Research activities in support to a European operational monitoring system for fossil CO2 emissions		
	LC-SPACE-14-TEC-2018-2019: Earth observation technologies		

Other topics related to Earth Observation (not including 2018-2020 topics above):

- Blue Growth – demonstrating an ocean of opportunities (H2020-BG-2016-2017):
  - BG-9-2016: An integrated Arctic observing system
  - BG-12-2016: Towards an integrated Mediterranean Sea Observing System
- Sustainable Food Security – resilient agri-food chains (H2020-SFS-2016-2017):
  - SFS-43-2017: Earth Observation services for the monitoring of agricultural production in Africa
- Climate Action, Environment, Resource Efficiency and Raw Materials - Earth Observation (H2020-SC5-2016-2017):
  - SC5-18-2017 - Novel in-situ observation systems o SC5-19-2017 - Coordination of citizens' observatories initiatives
  - SC5-20-2016 - European data hub of the GEOSS information system
- Competitiveness of the European Space Sector: Technology and Science (H2020- COMPET-2017)
  - COMPET-2-2017: Competitiveness in Earth observation mission technologies
- SME Instrument (H2020-SMEInst-2016-2017), although not dedicated uniquely to Earth Observation, is particularly well suited for SMEs addressing space based applications
  - SMEInst-04-2016-2017: Engaging SMEs in space research and development
  - SMEInst-12-2016-2017: Boosting the potential of small businesses in the areas and priorities of Societal Challenge 5

2. Ongoing and completed EO Big Data projects – ESA

Title	Description
Grid Enabled Service Support Environment	Developments of models and components to build data processing and analysis infrastructures. Integration with current and future vegetation processing chain.
Data Mining For Analysis and Exploitation of Next Generation of Time Series	General analytical methods for the exploitation of the information contained in Satellite Image Time Series Main focus on the information extraction in the form of “categories of evolution” and elaboration of technologies to classify the evolutions processes of observed scenes.

Open Source Image Retrieval - Integration of Developed tools	Develop, implement and integrate tools for Earth Observation (EO) Content Based Image Retrieval (CBIR) into a powerful and ready-to-use Open-Source platform. Definition and optimization of features guiding the search.
BIDS-RAF	Identify, evaluate and develop technologies and components to prototype a framework for a reference architecture for Big Data from Space

ANNEX VI – ABOUT BDVA

The **Big Data Value Association (BDVA**) is an industry-driven international non–for-profit organisation with over 190 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organizations.
The full list of BDVA members can be found here.

BDVA is the private counterpart to the EU Commission to implement the Big Data Value PPP program. The European Commission and Europe's data industry committed to invest €2.5 billion in this public-private partnership (PPP) that aims to strengthen the data sector and put Europe at the forefront of the global data race. The EU earmarked over €500 million of investment over 5 years (2016-2020) from Horizon 2020 for this purpose.
The **mission** of the BDVA is to develop the Innovation Ecosystem that will enable the data-driven digital transformation in Europe delivering maximum economic and societal benefit, and, achieving and sustaining Europe's leadership on **Big Data Value creation** and **Artificial Intelligence**.

To achieve this mission, the BDVA has defined in 2017 four strategic priorities:

- **Develop Data Innovation Recommendations**: Providing guidelines and recommendations on data innovation to the industry, researchers, markets and policy makers.
- **Develop Ecosystem**: Developing and Strengthening the European Big Data Value Ecosystem.
- **Guiding Standards**: Driving Big Data standardisation and interoperability priorities / influencing Standardisation bodies and industrial alliances.
- **Know-How and Skills**: Improve the adoption of Big Data through the exchange of knowledge, skills and best practices.

BDVA enables existing regional multi-partner cooperation, to collaborate at European level through the provision of tools and knowhow to support the co-creation, development and experimentation of pan-European data-driven applications and services, and knowhow exchange.

Since 2014 BDVA publishes a Big Data Value **Strategic Research and Innovation Agenda (SRIA)**. The SRIA v 4 has been launched in November 2017 and explains the strategic importance of Big Data, describes the data value chain and the central role of ecosystems, details a vision for Big Data Value in Europe in 2020, and sets out the objectives and goals to be accomplished by the PPP within the European research and innovation landscape of Horizon 2020 (H2020) and at both national and regional levels.
BDVA organizes its work in **Task Forces**, where its members engage and influence with the aim for BDVA to be the European Big Data Value reference point. The list of task forces and subgroups can be found on the BDVA website: http://bdva.eu/?q=task-force-overview.
BDVA organises an **annual conference** (European Big Data Value Forum, merge of the BDVA Summit and the European Data Forum) in collaboration with its members, Industry, research and the European Commission. Find the link to the latest edition here.
BDVA also organises regular meetings and workshops for members with involvement of many external public and private players (European Commission, Standardisation bodies, other industrial associations, etc).